

416 Rec'd PCT/PTO 12 JUN 2000

PCT/AU98/01023

WO 99/29870

1

*Porphyromonas gingivalis* polypeptides and nucleotides**FIELD OF THE INVENTION**

5        The present invention relates to *P. gingivalis* nucleotide sequences, *P. gingivalis* polypeptides and probes for detection of *P. gingivalis*. The *P. gingivalis* polypeptides and nucleotides can be used in compositions for use in raising an immune response in a subject against *P. gingivalis* and treating or preventing or reducing the severity of the condition known as  
10      periodontitis.

**BACKGROUND OF THE INVENTION**

15      Periodontal diseases are bacterial-associated inflammatory diseases of the supporting tissues of the teeth and range from the relatively mild form of gingivitis, the non-specific, reversible inflammation of gingival tissue to the more aggressive forms of periodontitis which are characterised by the destruction of the tooth's supporting structures. Periodontitis is associated with a subgingival infection of a consortium of specific Gram-negative  
20      bacteria that leads to the destruction of the periodontium and is a major public health problem. One bacterium that has attracted considerable interest is *P. gingivalis* as the recovery of this microorganism from adult periodontitis lesions can be up to 50% of the subgingival anaerobically cultivable flora, whereas *P. gingivalis* is rarely recovered, and then in low numbers, from healthy sites. A proportional increase in the level of  
25      *P. gingivalis* in subgingival plaque has been associated with an increased severity of periodontitis and eradication of the microorganism from the cultivable subgingival microbial population is accompanied by resolution of the disease. The progression of periodontitis lesions in non-human primates  
30      has been demonstrated with the subgingival implantation of *P. gingivalis*. These findings in both animals and humans suggest a major role for *P. gingivalis* in the development of adult periodontitis.

35      *P. gingivalis* is a black-pigmented, anaerobic, asaccharolytic, proteolytic Gram-negative rod that obtains energy from the metabolism of specific amino acids. The microorganism has an absolute growth requirement for iron, preferentially in the form of haeme or its Fe(III)

oxidation product haemin and when grown under conditions of excess haemin is highly virulent in experimental animals. A number of virulence factors have been implicated in the pathogenicity of *P. gingivalis* including the capsule, adhesins, cytotoxins and extracellular hydrolytic enzymes.

- 5        In order to develop an efficacious and safe vaccine to prevent, eliminate or reduce *P. gingivalis* colonisation it is necessary to identify and produce antigens that are involved in virulence that have utility as immunogens possibly through the generation of specific antibodies. Whilst it is possible to attempt to isolate antigens directly from cultures of *P. gingivalis*
- 10      this is often difficult. For example as mentioned above, *P. gingivalis* is a strict anaerobe and can be difficult to isolate and grow. It is also known that, for a number of organisms, when cultured *in vitro* that many virulence genes are down regulated and the encoded proteins are no longer expressed. If conventional chemistry techniques were applied to purify vaccine candidates
- 15      potentially important (protective) molecules may not be identified. With DNA sequencing, as the gene is present (but not transcribed) even when the organism is grown *in vitro* it can be identified, cloned and produced as a recombinant DNA protein. Similarly, a protective antigen or therapeutic target may be transiently expressed by the organism *in vitro* or produced in
- 20      low levels making the identification of these molecules extremely difficult by conventional methods.

- With serological identification of therapeutic targets one is limited to those responses which are detectable using standard methods such as Western Blotting or ELISA. The limitation here is the both the level of response that is generated by the animal or human and determining whether this response is protective, damaging or irrelevant. No such limitation is present with a sequencing approach to the identification of potential therapeutic or prophylactic targets.

- It is also well known that *P. gingivalis* produces a range of broadly active proteases (University of Melbourne International Patent Application No PCT/AU 96/00673, US Patent Nos 5,475,097 and 5,523,390), which make the identification of intact proteins difficult because of their degradation by these proteases.

**SUMMARY OF THE INVENTION**

The present inventors have attempted to isolate *P. gingivalis* nucleotide sequences which can be used for recombinant production of

- 5    *P. gingivalis* polypeptides and to develop nucleotide probes specific for *P. gingivalis*. The DNA sequences listed below have been selected from a large number of *P. gingivalis* sequences according to their indicative potential as vaccine candidates. This intuitive step involved comparison of the deduced protein sequence from the *P. gingivalis* DNA sequences to the
- 10   known protein sequence databases. Some of the characteristics used to select useful vaccine candidates include; the expected cellular location, such as outer membrane proteins or secreted proteins, particular functional activities of similar proteins such as those with an enzymatic or proteolytic activity, proteins involved in essential metabolic pathways that when
- 15   inactivated or blocked may be deleterious or lethal to the organism, proteins that might be expected to play a role in the pathogenesis of the organism eg. red cell lysis, cell agglutination or cell receptors and proteins which are paralogues to proteins with proven vaccine efficacy.

- In a first aspect the present invention consists an isolated antigenic
- 20   *Porphyromonas gingivalis* polypeptide, the polypeptide comprising; an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or an amino acid sequence at least 85%, preferably at least 95%,
- 25   identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532; or at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected
- 30   from the group consisting of SEQ. ID. NO. 265 to SEQ. ID. NO. 528, SEQ. ID. NO. 531 and SEQ. ID. NO. 532.
- In an embodiment of the present invention the polypeptide comprises;
- 35   an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532; or

an amino acid sequence at least 85%, preferably at least 95%, identical to an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532; or

5 at least 40 amino acids having a contiguous sequence of at least 40 amino acids identical to a contiguous amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532.

10 As used herein % identity for polypeptides is to be calculated using the alignment algorithm of Needleman and Munsch (9) using a standard protein scoring matrix (Blosum 50).

In a preferred embodiment of the present invention the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386, SEQ. ID. NO. 424, SEQ. ID. NO. 425, SEQ. ID. NO. 434, 15 SEQ. ID. NO. 447, SEQ. ID. NO. 458, SEQ. ID. NO. 475, SEQ. ID. NO. 498, SEQ. ID. NO. 499, SEQ. ID. NO. 500, SEQ. ID. NO. 501, SEQ. ID. NO. 387, SEQ. ID. NO. 400, SEQ. ID. NO. 411, SEQ. ID. NO. 419, SEQ. ID. NO. 420, SEQ. ID. NO. 427, SEQ. ID. NO. 429, SEQ. ID. NO. 433, SEQ. ID. NO. 437, 20 SEQ. ID. NO. 438, SEQ. ID. NO. 443, SEQ. ID. NO. 444, SEQ. ID. NO. 448, SEQ. ID. NO. 449, SEQ. ID. NO. 452, SEQ. ID. NO. 455, SEQ. ID. NO. 457, SEQ. ID. NO. 459, SEQ. ID. NO. 461, SEQ. ID. NO. 462, SEQ. ID. NO. 463, SEQ. ID. NO. 467, SEQ. ID. NO. 468, SEQ. ID. NO. 469, SEQ. ID. NO. 482, 25 SEQ. ID. NO. 484, SEQ. ID. NO. 485, SEQ. ID. NO. 494, SEQ. ID. NO. 508, SEQ. ID. NO. 509, SEQ. ID. NO. 510, SEQ. ID. NO. 520, SEQ. ID. NO. 521, SEQ. ID. NO. 522, SEQ. ID. NO. 525, SEQ. ID. NO. 526, SEQ. ID. NO. 528, 30 SEQ. ID. NO. 389, SEQ. ID. NO. 390 and SEQ. ID. NO. 391.

In another preferred embodiment of the present invention the polypeptide comprises an amino acid sequence selected from the group consisting of residue 422 to residue 531 of SEQ. ID. NO. 303, residue 534 to residue 582 of SEQ. ID. NO. 303, residue 127 to residue 232 of SEQ. ID. NO. 301, residue 240 to residue 259 of SEQ. ID. NO. 301, residue 139 to residue 156 of SEQ. ID. NO. 295, residue 160 to residue 178 of SEQ. ID. NO. 295, residue 180 to residue 207 of SEQ. ID. NO. 295, residue 221 to residue 257 of SEQ. ID. NO. 295, residue 259 to residue 323 of SEQ. ID. NO. 295, residue 35 885 to residue 985 of SEQ. ID. NO. 299, residue 147 to residue 259 of SEQ. ID. NO. 363, residue 140 to residue 252 of SEQ. ID. NO. 344, residue 247 to

residue 356 of SEQ. ID. NO. 353, residue 359 to residue 391 of SEQ. ID. NO. 353, residue 120 to residue 254 of SEQ. ID. NO. 300, residue 287 to residue 311 of SEQ. ID. NO. 286, residue 313 to residue 352 of SEQ. ID. NO. 286, residue 354 to residue 401 of SEQ. ID. NO. 286, residue 208 to residue 252 of SEQ. ID. NO. 287, residue 259 to residue 373 of SEQ. ID. NO. 287, residue 5 to residue 120 of SEQ. ID. NO. 293, residue 123 to residue 139 of SEQ. ID. NO. 293, residue 233 to residue 339 of SEQ. ID. NO. 265, residue 67 to residue 228 of SEQ. ID. NO. 278, residue 130 to residue 172 of SEQ. ID. NO. 274, residue 174 to residue 238 of SEQ. ID. NO. 274, residue 99 to residue 112 of SEQ. ID. NO. 274, residue 114 to residue 128 of SEQ. ID. NO. 274, residue 26 to residue 69 of SEQ. ID. NO. 285, residue 71 to residue 128 of SEQ. ID. NO. 285, residue 130 to residue 146 of SEQ. ID. NO. 285, residue 620 to residue 636 of SEQ. ID. NO. 327, residue 638 to residue 775 of SEQ. ID. NO. 327, residue 397 to residue 505 of SEQ. ID. NO. 301, residue 528 to residue 545 of SEQ. ID. NO. 301, residue 556 to residue 612 of SEQ. ID. NO. 301, residue 614 to residue 631 of SEQ. ID. NO. 301, residue 633 to residue 650 of SEQ. ID. NO. 301, residue 553 to residue 687 of SEQ. ID. NO. 299, residue 305 to residue 447 of SEQ. ID. NO. 289, residue 1 to residue 52 of SEQ. ID. NO. 364, residue 65 to residue 74 of SEQ. ID. NO. 364, residue 486 to residue 604 of SEQ. ID. NO. 275, residue 158 to residue 267 of SEQ. ID. NO. 272, residue 270 to residue 282 of SEQ. ID. NO. 272, residue 163 to residue 237 of SEQ. ID. NO. 273, residue 240 to residue 251 of SEQ. ID. NO. 273, residue 213 to residue 344 of SEQ. ID. NO. 282, residue 183 to residue 324 of SEQ. ID. NO. 292, residue 327 to residue 341 of SEQ. ID. NO. 292, residue 352 to residue 372 of SEQ. ID. NO. 292, residue 141 to residue 166 of SEQ. ID. NO. 271, residue 168 to residue 232 of SEQ. ID. NO. 271, residue 1 to residue 13 of SEQ. ID. NO. 302, residue 15 to residue 28 of SEQ. ID. NO. 302, residue 30 to residue 72 of SEQ. ID. NO. 302, residue 476 to residue 529 of SEQ. ID. NO. 277, residue 41 to residue 146 of SEQ. ID. NO. 299, residue 149 to residue 162 of SEQ. ID. NO. 299, residue 166 to residue 177 of SEQ. ID. NO. 299, residue 192 to residue 203 of SEQ. ID. NO. 299, residue 71 to residue 343 of SEQ. ID. NO. 290, residue 346 to residue 363 of SEQ. ID. NO. 290, residue 36 to residue 240 of SEQ. ID. NO. 331, residue 242 to residue 270 of SEQ. ID. NO. 331, residue 1 to residue 192 of SEQ. ID. NO. 375, residue 266 to residue 290 of SEQ. ID. NO. 375, residue 23 to residue 216 of SEQ. ID. NO. 279, residue 220 to residue 270 of SEQ. ID. NO. 279, residue

285 to residue 386 of SEQ. ID. NO. 279, residue 84 to residue 234 of SEQ. ID. NO. 297, residue 248 to residue 259 of SEQ. ID. NO. 297, residue 261 to residue 269 of SEQ. ID. NO. 297, residue 275 to residue 402 of SEQ. ID. NO. 294, residue 1 to residue 171 of SEQ. ID. NO. 298, residue 403 to residue 417 of SEQ. ID. NO. 307, residue 420 to residue 453 of SEQ. ID. NO. 307, residue 456 to residue 464 of SEQ. ID. NO. 307, residue 468 to residue 690 of SEQ. ID. NO. 307, residue 1 to residue 285 of SEQ. ID. NO. 304, residue 287 to residue 315 of SEQ. ID. NO. 304, residue 318 to residue 336 of SEQ. ID. NO. 304, residue 255 to residue 269 of SEQ. ID. NO. 342, residue 271 to residue 337 of SEQ. ID. NO. 342, residue 347 to residue 467 of SEQ. ID. NO. 281, residue 116 to residue 136 of SEQ. ID. NO. 375, residue 138 to residue 357 of SEQ. ID. NO. 375, residue 133 to residue 423 of SEQ. ID. NO. 364, residue 141 to residue 299 of SEQ. ID. NO. 305, residue 202 to residue 365 of SEQ. ID. NO. 296, residue 134 to residue 426 of SEQ. ID. NO. 288, residue 1 to residue 218 of SEQ. ID. NO. 276, residue 1 to residue 246 of SEQ. ID. NO. 280, residue 444 to residue 608 of SEQ. ID. NO. 364, residue 10 to residue 686 of SEQ. ID. NO. 283, residue 1 to residue 148 of SEQ. ID. NO. 296, residue 1 to residue 191 of SEQ. ID. NO. 287, residue 193 to residue 204 of SEQ. ID. NO. 287, residue 209 to residue 373 of SEQ. ID. NO. 287, residue 211 to residue 470 of SEQ. ID. NO. 284, residue 472 to residue 482 of SEQ. ID. NO. 284, residue 133 to residue 144 of SEQ. ID. NO. 281, residue 146 to residue 336 of SEQ. ID. NO. 281, residue 1 to residue 264 of SEQ. ID. NO. 303, residue 265 to residue 295 of SEQ. ID. NO. 303, residue 297 to residue 326 of SEQ. ID. NO. 303, residue 328 to residue 338 of SEQ. ID. NO. 303, residue 247 to residue 356 of SEQ. ID. NO. 353, residue 358 to residue 391 of SEQ. ID. NO. 353, residue 257 to residue 288 of SEQ. ID. NO. 298, residue 290 to residue 385 of SEQ. ID. NO. 298, residue 245 to residue 256 of SEQ. ID. NO. 298, residue 422 to residue 802 of SEQ. ID. NO. 303, residue 803 to residue 814 of SEQ. ID. NO. 303, residue 139 to residue 156 of SEQ. ID. NO. 295, residue 160 to residue 340 of SEQ. ID. NO. 295, residue 145 to residue 361 of SEQ. ID. NO. 282, residue 363 to residue 387 of SEQ. ID. NO. 282, residue 398 to residue 471 of SEQ. ID. NO. 282, residue 573 to residue 679 of SEQ. ID. NO. 320, residue 27 to residue 168 of SEQ. ID. NO. 291, residue 170 to residue 183 of SEQ. ID. NO. 291, residue 185 to residue 415 of SEQ. ID. NO. 291, residue 1 to residue 301 of SEQ. ID. NO. 364, residue 114 to residue 702 of SEQ. ID. NO. 337, residue 377 to residue 412 of SEQ. ID. NO. 321,

residue 413 to residue 772 of SEQ. ID. NO. 321, residue 14 to residue 454 of SEQ. ID. NO. 265, residue 129 to residue 614 of SEQ. ID. NO. 268, residue 1 to residue 930 of SEQ. ID. NO. 300, residue 932 to residue 1046 of SEQ. ID. NO. 300, residue 1 to residue 301 of SEQ. ID. NO. 364, residue 1 to residue 42 of SEQ. ID. NO. 381, residue 44 to residue 973 of SEQ. ID. NO. 381, residue 1 to residue 93 of SEQ. ID. NO. 358, residue 95 to residue 179 of SEQ. ID. NO. 358, residue 181 to residue 227 of SEQ. ID. NO. 358, residue 114 to residue 702 of SEQ. ID. NO. 337, residue 1 to residue 659 of SEQ. ID. NO. 355, residue 661 to residue 907 of SEQ. ID. NO. 355, residue 1 to residue 131 of SEQ. ID. NO. 370, residue 133 to residue 601 of SEQ. ID. NO. 370, residue 1 to residue 813 of SEQ. ID. NO. 344, residue 377 to residue 412 of SEQ. ID. NO. 321, residue 413 to residue 772 of SEQ. ID. NO. 321, and residue 189 to residue 614 of SEQ. ID. NO. 364.

In a second aspect the present invention consists in an isolated antigenic *Porphyromonas gingivalis* polypeptide, the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ. ID. NO. 386 to SEQ. ID. NO. 528 and SEQ. ID. NO. 532 less the leader sequence set out in Table 3.

In a third aspect the present invention consists in an isolated DNA molecule, the DNA molecule comprising a nucleotide sequence which encodes the polypeptide of the first aspect the present invention or a sequence which hybridises thereto under stringent conditions.

It is preferred that the isolated DNA molecule comprises a nucleotide sequence selected from the group consisting of SEQ. ID. NO. 1 to SEQ. ID. NO. 264, SEQ. ID. NO. 529 and SEQ. ID. NO. 530.

In a fourth aspect the present invention consists in a recombinant expression vector comprising the DNA molecule of the second aspect of the present invention operably linked to a transcription regulatory element.

The present invention also provides a cell comprising this recombinant expression vector.

In a further aspect the present invention consists in a method for producing a *P. gingivalis* polypeptide comprising culturing the cell under conditions that permit expression of the polypeptide.

In yet a further aspect the present invention provides a composition for use in raising an immune response directed against *P. gingivalis* in a subject, the composition comprising an effective amount of at least one

polypeptide of the first aspect of the present invention, or at least one DNA molecule of the second aspect of the present invention, or both, and a pharmaceutically acceptable carrier. It is preferred that the pharmaceutically acceptable carrier is an adjuvant. In other aspects the present invention 5 provides methods of treating *P. gingivalis* infection in subject comprising the administration of the composition to the subject such that treatment of *P. gingivalis* infection occurs. The treatment may be prophylactic or therapeutic.

In yet another aspect the present invention provides an antibody 10 raised against a polypeptide of the first aspect the invention. The antibody may be polyclonal or monoclonal. The present invention also provides compositions including these antibodies. It is preferred that these compositions are adapted for oral use and may be, for example, dentrifrices, mouthwashes, etc.

15 In a still further aspect the present invention provides a nucleotide probe comprising at least 18 nucleotides and having a contiguous sequence of at least 18 nucleotides identical to a contiguous nucleotide sequence selected from the group consisting of SEQ. ID. NO. 1 to SEQ. ID. NO. 121, SEQ. ID. NO. 529, and sequences complementary thereto. It is preferred that 20 the probe further comprises a detectable label.

The present invention also provides a method for detecting the presence of *P. gingivalis* nucleic acid in a sample comprising:

- (a) contacting a sample with the nucleotide probe under conditions in which a hybrid can form between the probe and a *P. gingivalis* 25 nucleic acid in the sample; and  
(b) detecting the hybrid formed in step (a), wherein detection of a hybrid indicates the presence of a *P. gingivalis* nucleic acid in the sample.

### 30 DETAILED DESCRIPTION

#### Definitions

A purified or isolated polypeptide or a substantially pure preparation of a polypeptide are used interchangeably herein and, as used herein, mean a 35 polypeptide that has been separated from other proteins, lipids, and nucleic acids with which it naturally occurs. Preferably, the polypeptide is also

separated from substances, e.g., antibodies or gel matrix, e.g., polyacrylamide, which are used to purify it. Preferably, the polypeptide constitutes at least 10, 20, 50 70, 80 or 95% dry weight of the purified preparation. Preferably, the preparation contains: sufficient polypeptide to allow protein sequencing; at least 1, 10, or 100 mg of the polypeptide.

5 A purified preparation of cells refers to, in the case of plant or animal cells, an *in vitro* preparation of cells and not an entire intact plant or animal. In the case of cultured cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

10 A purified or isolated or a substantially pure nucleic acid, e.g., a substantially pure DNA, (are terms used interchangeably herein) is a nucleic acid which is one or both of the following: not immediately contiguous with both of the coding sequences with which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the naturally occurring genome of the organism from which the nucleic acid is derived; or which is substantially free of a nucleic acid with which it occurs in the organism from which the nucleic acid is derived. The term includes, for example, a recombinant DNA which is incorporated into a vector, e.g., into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other DNA sequences. Substantially pure DNA also includes a recombinant DNA which is part of a hybrid gene encoding additional *P. gingivalis* DNA sequence.

15 20 25 A "contig" as used herein is a nucleic acid representing a continuous stretch of genomic sequence of an organism.

An "open reading frame", also referred to herein as ORF, is a region of nucleic acid which encodes a polypeptide. This region may represent a portion of a coding sequence or a total sequence and can be determined from a stop to stop codon or from a start to stop codon.

30 35 As used herein, a "coding sequence" is a nucleic acid which is transcribed into messenger RNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the five prime terminus and a translation stop code at the three prime terminus. A

coding sequence can include but is not limited to messenger RNA synthetic DNA, and recombinant nucleic acid sequences.

- 5 A "complement" of a nucleic acid as used herein refers to an anti-parallel or antisense sequence that participates in Watson-Crick base-pairing with the original sequence.

A "gene product" is a protein or structural RNA which is specifically encoded by a gene.

- As used herein, the term "probe" refers to a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest.  
 10 Probes are often associated with or capable of associating with a label. A label is a chemical moiety capable of detection. Typical labels comprise dyes, radioisotopes, luminescent and chemiluminescent moieties, fluorophores, enzymes, precipitating agents, amplification sequences, and the like. Similarly, a nucleic acid, peptide or other chemical entity which  
 15 specifically binds to a molecule of interest and immobilizes such molecule is referred herein as a "capture ligand". Capture ligands are typically associated with or capable of associating with a support such as nitro-cellulose, glass, nylon membranes, beads, particles and the like. The specificity of hybridization is dependent on conditions such as the base pair composition of the nucleotides, and the temperature and salt concentration of the reaction. These conditions are readily discernible to one of ordinary skill in the art using routine experimentation.

- Homologous refers to the sequence similarity or sequence identity between two polypeptides or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two  
 25 sequences divided by the number of positions compared x 100.

The terms peptides, proteins, and polypeptides are used interchangeably herein.

- An "immunogenic component" as used herein is a moiety, such as an *P. gingivalis* polypeptide, analog or fragment thereof, that is capable of eliciting a humoral and/or cellular immune response in a host animal.  
 30  
 35

An "antigenic component" as used herein is a moiety, such as *P. gingivalis* polypeptide, analog or fragment thereof, that is capable of binding to a specific antibody with sufficiently high affinity to form a detectable antigen-antibody complex.

- 5 As used herein, the term "cell-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue. The term also covers so-called "leaky" promoters, which regulate expression of a selected
- 10 DNA primarily in one tissue, but cause expression in other tissues as well.

As used herein, the term "control sequence" refers to a nucleic acid having a base sequence which is recognized by the host organism to effect the expression of encoded sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include a promoter, ribosomal binding site, terminators, and in some cases operators; in eukaryotes, generally such control sequences include promoters, terminators and in some instances, enhancers. The term control sequence is intended to include at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is advantageous, for example, leader sequences.

- 20
- 25 As used herein, the term "operably linked" refers to sequences joined or ligated to function in their intended manner. For example, a control sequence is operably linked to coding sequence by ligation in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequence and host cell.

A "sample" as used herein refers to a biological sample, such as, for example, tissue or fluid isolated from an individual (including without limitation plasma, serum, cerebrospinal fluid, lymph, tears, saliva and tissue sections) or from *in vitro* cell culture constituents, as well as samples from the environment.

- 30
- 35 The practice of the invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, A Practical Guide to Molecular

Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, Volumes 1 and 2, IRL Press (1991), D.M. Glover and B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996), and F.M. Ausubel et al. (Editors), Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience (1988, including all updates until present). The disclosure of these texts are incorporated herein by reference.

10    Pharmaceutically Acceptable Carriers

The antibodies, polypeptides and DNA of the present invention can be included in compositions which include a carrier or diluent. These compositions include pharmaceutical compositions where the carrier or diluent will be pharmaceutically acceptable. Pharmaceutically acceptable carriers or diluents include those used in compositions suitable for oral, rectal, nasal, topical (including buccal and sublingual), vaginal, parenteral (including subcutaneous, intramuscular, intravenous, intradermal, intrathecal and epidural) administration. They are non-toxic to recipients at the dosages and concentrations employed. Representative examples of pharmaceutically acceptable carriers or diluents include, but are not limited to; water, isotonic solutions which are preferably buffered at a physiological pH (such as phosphate-buffered saline or Tris-buffered saline) and can also contain one or more of, mannitol, lactose, trehalose, dextrose, glycerol, ethanol or polypeptides (such as human serum albumin). The compositions may conveniently be presented in unit dosage form and may be prepared by any of the methods well known in the art of pharmacy.

As will be well understood by those skilled in the art alterations may be made to the amino acid sequences set out in the Sequence Listings. These alterations may be deletions, insertions, or substitutions of amino acid residues. The altered polypeptides can be either naturally occurring (that is to say, purified or isolated from a natural source) or synthetic (for example, by performing site-directed metagenesis on the encoding DNA). It is intended that such altered polypeptides which have at least 85%, preferably at least 95% identity with the sequences set out in the Sequence Listing are

within the scope of the present invention. Antibodies raised against these altered polypeptides will also bind to the polypeptides having one of the sequences set out in the Sequence Listings. The level of % identity is to be calculated as set out above.

- 5      Protein sequences are homologous if they are related by divergence from a common ancestor. Consequently, a species homologue of the protein will be the equivalent protein which occurs naturally in another species. Within any one species a homologue may exist as numerous allelic variants, and these will be considered homologues of the protein. Allelic variants and  
10     species homologues can be obtained by following standard techniques known to those skilled in the art.

An allelic variant will be a variant that is naturally occurring within an individual organism.

15     **Mutants, Variants and Homology - Nucleic Acids**

- Mutant polynucleotides will possess one or more mutations which are deletions, insertions, or substitutions of nucleotide residues. Mutants can be either naturally occurring (that is to say, isolated from a natural source) or synthetic (for example, by performing site-directed mutagenesis on the DNA). It is thus apparent that polynucleotides of the invention can be either naturally occurring or recombinant (that is to say prepared using recombinant DNA techniques).

An allelic variant will be a variant that is naturally occurring within an individual organism.

- 25     Nucleotide sequences are homologous if they are related by divergence from a common ancestor. Consequently, a species homologue of the polynucleotide will be the equivalent polynucleotide which occurs naturally in another species. Within any one species a homologue may exist as numerous allelic variants, and these will be considered homologues of the  
30     polynucleotide. Allelic variants and species homologues can be obtained by following standard techniques known to those skilled in the art.

**Antibody Production**

- 35     Antibodies, either polyclonal or monoclonal, which are specific for a polypeptide of the present invention can be produced by a person skilled in the art using standard techniques such as, but not limited to, those described

by Harlow *et al.* Antibodies: A Laboratory Manual, Cold Springs Harbor Laboratory Press (1988), and D. Catty (editor), Antibodies: A Practical Approach, IRL Press (1988).

- Various procedures known in the art may be used for the production of polyclonal antibodies to epitopes of a protein. For the production of polyclonal antibodies, a number of host animals are acceptable for the generation of antibodies by immunization with one or more injections of a polypeptide preparation, including but not limited to rabbits, mice, rats, etc. Various adjuvants may be used to increase the immunological response in the host animal, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminium hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*.

- A monoclonal antibody to an epitope of a protein may be prepared by using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Kohler and Milstein (1975, *Nature* 256, 493-497), and the more recent human B-cell hybridoma technique (Kesber *et al.* 1983, *Immunology Today* 4:72) and EBV-hybridoma technique (Cole *et al.* 1985, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). In addition, techniques developed for the production of "chimeric antibodies" by splicing the genes from antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity may be used (Morrison *et al.* 1984, *Proc. Natl. Acad. Sci.*, 81:6851-6855; Neuberger *et al.* 1984 *Nature* 312:604-608; Takeda *et al.* 1985 *Nature* 31:452-454). Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce 4-specific single chain antibodies.

- Recombinant human or humanized versions of monoclonal antibodies are a preferred embodiment for human therapeutic applications. Humanized antibodies may be prepared according to procedures in the literature (e.g. Jones *et al.* 1986, *Nature* 321:522-25; Reichman *et al.* 1988 *Nature* 332:323-27; Verhoeven *et al.* 1988, *Science* 239:1534-36). The

recently described "gene conversion metagenesis" strategy for the production of humanized monoclonal antibody may also be employed in the production of humanized antibodies (Carter et al. 1992 Proc. Natl. Acad. Sci. U.S.A. 89:4285-89). Alternatively, techniques for generating the recombinant phase

- 5 library of random combinations of heavy and light regions may be used to prepare recombinant antibodies (e.g. Huse et al. 1989 Science 246:1275-81).

Antibody fragments which contain the idiotype of the molecule such as Fu F(ab1) and F(ab2) may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab) E2

- 10 fragment which can be produced by pepsin digestion of the intact antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragment, and the two Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent. Alternatively, Fab expression libraries may be constructed (Huse et al. 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragment with the desired specificity to a protein.

#### Adjuvants

- 20 "Adjuvant" means a composition comprised of one or more substances that enhances the immunogenicity and efficacy of a vaccine composition. Non-limiting examples of suitable adjuvants include squalane and squalene (or other oils of animal origin); block copolymers; detergents such as Tween®-80; Quil® A, mineral oils such as Drakeol or Marcol, vegetable oils such as peanut oil; Corynebacterium-derived adjuvants such as Corynebacterium parvum; Propionibacterium-derived adjuvants such as Propionibacterium acne; Mycobacterium bovis (Bacillus Calmette and Guérin or BCG); interleukins such as interleukin 2 and interleukin-12; monokines such as interleukin 1; tumour necrosis factor; interferons such as gamma interferon; combinations such as saponin-aluminium hydroxide or Quil-A aluminium hydroxide; liposomes; ISCOM adjuvant; mycobacterial cell wall extract; synthetic glycopeptides such as muramyl dipeptides or other derivatives; Avridine; Lipid A; dextran sulfate; DEAE-Dextran or DHAEDextran with aluminium phosphate; carboxypolymethylene such as Carbopol® EMA; acrylic copolymer emulsions such as Neocryl A640 (e.g. U.S.

Pat. No. 5,047,238); vaccinia or animal poxvirus proteins; sub-viral particle adjuvants such as cholera toxin, or mixtures thereof.

- As used herein, stringent conditions are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% NaDODSO<sub>4</sub> at 50°C; (2) employ during hybridisation a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS and 10% dextran sulfate at 42°C in 0.2 x SSC and 0.1% SDS

15

As will be understood the present invention includes within its scope DNA vaccination. Further information regarding DNA vaccination may be found in Donnelly et al, Journal of Immunological Methods 176(1994) 145-152, the disclosure of which is incorporated herein by reference.

20

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer, or group of elements or integers.

25

### **Preparation of the *P. gingivalis* library for sequencing.**

- To determine the DNA sequence of *P. gingivalis* genomic DNA was isolated from *P. gingivalis* strain W50 (ATCC 53978) essentially by the method described by Mamur J. (J. Mol. Biol. 3, 208-218, 1961). Cloning of DNA fragments was performed essentially as described by Fleischmann et al., (Science; 269, 496-512, 1995)(2). Briefly, purified genomic DNA from *P. gingivalis* was nebulized to fragment the DNA and was treated with Bal31 nuclease to create blunt ends then run twice through preparative 1% agarose gels. DNA fragments of 1.6-2.0 kb were excised from the gel and the DNA recovered. This DNA was then ligated to the vector pUC18 (*Sma*I digested and dephosphorylated; Pharmacia) and electrophoresed through a 1%

35

preparative agarose gel. The fragment comprising linear vector plus one insert was excised, purified and this process repeated to reduce any vector without insert contamination. The recovered vector plus insert DNA was blunt-ended with T4 DNA polymerase, then a final ligation to produce 5 circular DNA was performed. Aliquots of Epicurian Coli Electroporation-Competent Cells (Stratagene) were transformed with the ligated DNA and plated out on SOB agar antibiotic diffusion plates containing X-gal and incubated at 37°C overnight. Colonies with inserts appeared white and those without inserts (vector alone) appeared blue. Plates were stored at 4°C until 10 the white clones were picked and expanded for the extraction of plasmid DNA for sequencing.

#### DNA sequencing

Plasmid DNA was prepared by picking bacterial colonies into 1.5ml 15 of LB, TB or SOB broth supplemented with 50-100ug/ml Ampicillin in 96 deep well plates. Plasmid DNA was isolated using the QIAprep Spin or QIAprep 96 Turbo miniprep kits (QIAGEN GmbH, Germany). DNA was eluted into a 96 well gridded array and stored at -20C.

Sequencing reactions were performed using ABI PRISM Dye 20 Terminator and ABI PRISM BIGDye Terminator Cycle Sequencing Ready Reaction kits with AmpliTaq DNA polymerase FS (PE Applied Biosystems, Foster City, CA) using the M13 Universal forward and reverse sequencing primers. Sequence reactions were conducted on either a Perkin-Elmer GeneAmp 9700 (PE Applied Biosystems) or Hybaid PCR Express (Hybaid, 25 UK) thermal cyclers. Sequencing reactions were analysed on ABI PRISM 377 DNA sequencers (PE Applied Biosystems).

The sequences obtained are set out below. The relationship between these sequences is set out in Table 1. The initiation codon was calculated using a combination of sequence homology alignment (FASTA), signal 30 sequence prediction (PSORT, SignalP) or ORF prediction (GeneMark).

**Table 1:** Reference table indicating the relationships of each sequence ID to the selected proteins.

Protein name	DNA sequence of complete ORF	Amino acid sequence of complete ORF	DNA sequence of protein	Amino acid sequence of protein
PG1	1	265	122	386
PG10	2	266	123	387
PG100	3	267	124	388
PG101	4	268		
PG102	5	269	125, 126	389, 390
PG104	6	270	127	391
PG105	7	271	128	392
PG106	8	272	129	393
PG107	9	273	130, 131, 132	394, 395, 396
PG108	10	274	133	397
PG109	11	275	134, 135	398, 399
PG11	12	276	136	400
PG110	13	277	137	401
PG111	14	278		
PG112	15	279	138, 139	402, 403
PG113	16	280	140	404
PG114	17	281	141	405
PG115	18	282	142	406
PG116	19	283	143	407
PG117	20	284	144	408
PG118	21	285	145	409
PG119	22	286	146	410
PG12	23	287	147	411
PG120	24	288	148	412

Protein name	DNA sequence of complete ORF	Amino acid sequence of complete ORF	DNA sequence of protein	Amino acid sequence of protein
PG121	25	289	149	413
PG122	26	290	150	414
PG123	27	291	151	415
PG124	28	292	152	416
PG125	29	293	153	417
PG126	30	294	154	418
PG13	31	295	155	419
PG14	32	296	156	420
PG15	33	297	157	421
PG16	34	298	158	422
PG18	35	299	159	423
PG2	36	300	160, 161	424, 425
PG21	37	301	162	426
PG22	38	302	163	427
PG23	39	303	164	428
PG24	40	304	165	429
PG25	41	305	166	430
PG27	42	306	167	431
PG28	43	307	168	432
PG29	44	308	169	433
PG3	45	309	170	434
PG30	46	310	171	435
PG31	47	311	172	436
PG32	48	312	173	437
PG33	49	313	174	438
PG34	50	314	175, 176	439, 440
PG35	51	315	177	441
PG36	52	316	178	442
PG37	53	317	179, 180	443, 444
PG38	54	318	181	445

Protein name	DNA sequence of complete ORF	Amino acid sequence of complete ORF	DNA sequence of protein	Amino acid sequence of protein
PG39	55	319	182	446
PG4	56	320	183	447
PG40	57	321	184	448
PG41	58	322	185	449
PG42	59	323	186	450
PG43	60	324	187	451
PG44	61	325	188	452
PG45	62	326	189	453
PG46	63	327	190	454
PG47	64	328	191	455
PG48	65	329	192	456
PG49	66	330	193	457
PG5	67	331	194	458
PG50	68	332	195	459
PG51	69	333	196	460
PG52	70	334	197	461
PG53	71	335	198	462
PG54	72	336	199	463
PG55	73	337	200	464
PG56	74	338	201, 202	465, 466
PG57	75	339	203, 204, 205	467, 468, 469
PG58	76	340	206, 207	470, 471
PG59	77	341	208, 209, 210	472, 473, 474
PG6	78	342	211	475
PG60	79	343	212	476
PG61	80	344	213	477
PG62	81	345	214	478
PG63	82	346	215	479
PG64	83	347	216	480
PG65	84	348	217	481

Protein name	DNA sequence of complete ORF	Amino acid sequence of complete ORF	DNA sequence of protein	Amino acid sequence of protein
PG66	85	349	218	482
PG67	86	350	219	483
PG68	87	351	220, 221	484, 485
PG69	88	352	222	486
PG7	89	353	223	487
PG70	90	354	224	488
PG71	91	355	225	489
PG72	92	356	226	490
PG73	93	357	227	491
PG74	94	358	228	492
PG75	95	359	229	493
PG76	96	360	230	494
PG77	97	361	231	495
PG78	98	362	232	496
PG79	99	363	233	497
PG8	100	364	234, 235, 236, 237	498, 499, 500, 501
PG80	101	365	238	502
PG81	102	366	102	366
PG82	103	367	239	503
PG83	104	368	240	504
PG84	105	369	241, 242	505, 506
PG85	106	370	243	507
PG86	107	371	244, 245	508, 509
PG87	108	372	246	510
PG88	109	373	247, 248, 249	511, 512, 513
PG89	110	374	250	514
PG9	111	375	251, 252, 253	515, 516, 517
PG90	112	376	254, 255	518, 519
PG91	113	377	256	520

Protein name	DNA sequence of complete ORF	Amino acid sequence of complete ORF	DNA sequence of protein	Amino acid sequence of protein
PG92	114	378	257	521
PG93	115	379	258	522
PG94	116	380	259	523
PG95	117	381	260	524
PG96	118	382	261	525
PG97	119	383	262	526
PG98	120	384	263	527
PG99	121	385	264	528
PG127	529	531	530	532

**DNA sequence analysis**

5

DNA files in FASTA format were converted to GCG format files and imported into a database. The DNA files were translated into amino acid files using the program Flip obtained from ANGIS(Australian Genomic Information Service, University of Sydney, Australia).

10 A series of bioinformatic analyses were performed on the proteins in order to select potential vaccine candidates. The programs used were FASTA homology searching (1), PSORT (2,3), SignalP (4), TopPred (5), and GeneMark (6). The proteins and their bioinformatic results were stored in the custom written database for search and retrieval of proteins with the desired characteristics

15 The FASTA homology results for these proteins were then examined for any alignment with a protein suggesting surface location or vaccine efficacy. All proteins were searched for homology against a non-redundant bacterial protein database compiled by ANGIS using the FASTA algorithm.

The settings used for the FASTA searches were Ktup = 2, gap creation 20 penalty = -12, gap extension penalty = -2, width for deriving alignment in opt = 16 and the Blosum 50 scoring matrix. Individual FASTA search results were examined for significant homology by statistical probability and amino acid alignments. The results are set out in Table 2.

Protein files were then trimmed to the first, second, third, fourth and fifth methionine residues using a protein trimming program (ANGIS). The trimmed proteins were then subjected to PSORT analysis for the detection of signal sequences and the prediction of cell location. Proteins exhibiting a  
5 PSORT probability of outer membrane >0.8 were considered to indicate surface localisation. A second signal sequence detection program SignalP was also performed and, in certain instances, this program detected signals not identified with PSORT. All proteins identified by other methods were also analysed by PSORT and SignalP. Previously, the C-terminal amino acid  
10 of bacterial outer membrane proteins has been shown to be important for the assembly of the protein on the outer membrane (7). A typical structure definition for outer membrane proteins has been determined as the presence of a signal sequence at the N-terminus and a tyrosine or phenylalanine at the C-terminus. A number of the selected proteins exhibit this characteristic  
15 structure. The program TopPred was used to determine the presence and number of membrane spanning domains (MSDs) and the presence of such sequences indicates a preference to be attached to membranes such as the outer membrane. The results of PSORT, SignalP and TopPred analyses with the C-terminal amino acids of the selected proteins are set out in Table 3.

20 The 70 amino acids from the C-terminus of a number of *P. gingivalis* outer membrane proteins share 50-100% protein sequence identity. These proteins included RGP1, RGP2, KGP, HagA, HagC, HagD, prtH and prtT. This conserved motif may be involved in the attachment or sorting of proteins to the outer membrane. The protein data set was searched using FASTA  
25 homology as described above and a number of novel proteins were identified which demonstrate similar motifs at their C-termini. The results are listed in Table 4

30 The TonBIII box is a 30 amino acid motif present within TonB outer membrane receptors in a wide variety of bacteria. The TonBIII box of *P. gingivalis* (8) was used to search the protein data set for homology by FASTA as described above. Those proteins demonstrating significant homology are listed in Table 5.

24

Table 2: FASTA protein homology results of complete ORFs against a non-redundant protein database.

Protein name	Homology description	Genbank accession number	Length of homolog protein	FASTA homology results		
				Length of <i>P. gingivalis</i> protein	Identity %	Overlap E value
PG1	48kD outer membrane protein, <i>Actinobacillus pleuropneumoniae</i>	U24492	449aa	451aa	32	454aa 1.40E-42
PG2	Outer membrane protein (susC), <i>Bacteroides thetaiotaomicron</i>	L49338	1098aa	1017aa, 1014aa	28	1099aa 4.00E-32
PG3	Outer membrane porin F adhesin, <i>Pseudomonas fluorescens</i>	U19743	317aa	223aa	35	187aa 1.10E-10
PG4	Outer membrane protein A, <i>Escherichia fergusonii</i>	M63352	243aa	672aa	48	88aa 4.10E-10
PG5	Adhesin protein (AdcA), <i>Streptococcus pneumoniae</i>	Z71552	423aa	315aa	25	279aa 9.40E-15
PG6	Hemolysin A (phyA), <i>Prevotella melaninogenica</i>	U27587	332aa	324aa	60	300aa 3.00E-74
PG7	Hemolysin (lycC), <i>Serpulina hydrolysans</i>	X73141	268aa	404aa	33	268aa 1.40E-24
PG8	Heme uptake protein A, <i>Bacteroides fragilis</i>	X97122	431aa	598aa, 550aa, 456aa,	79	417aa 6.70E-121
PG9	Internalin A (InlA), <i>Lysteria monocytogenes</i>	M67471	744aa	1260aa, 1232aa, 1174aa	38	340aa 7.30E-23

25

Protein name	Homology description	Genbank accession number	Length of homolog protein	Length of P. gingivalis protein	FASTA homology results		
					Identity %	Overlap	E value
PG10	Macrophage infectivity potentiator (MIP), Legionella oakridensis.	U92214	234aa	185aa	50	201aa	4.70E-31
PG11	Haemagglutinin (phg), Prevotella intermedia	AF017417	309aa	313aa	44	309aa	3.60E-44
PG12	Outer membrane lipoprotein, Haemophilus influenzae	M68502	274aa	271aa	36	254aa	9.60E-27
PG13	Ferric receptor (cfa), Campylobacter coli	U80812	696aa	757aa	24	625aa	1.20E-18
PG14	36KD antigen, Helicobacter pylori	U86610	329aa	331aa	37	326aa	1.10E-35
PG15	Outer membrane protein, Escherichia amylovora	X77921	377aa	267aa	30	253aa	5.40E-08
PG16	C terminal protease, Bartonella bacilliformis	J37094	434aa	569aa	38	357aa	3.00E-35
PG18	Protein-export membrane protein (secD), Helicobacter pylori	AE000652	503aa	961aa	32	811aa	1.10E-36
PG21	Surface antigen gene, Methanotarsina mezel	X34710	783aa	821aa	37	331aa	6.20E-33
PG22	Alpha-hemolysin gene, Aeromonas hydrophila	L36482	85aa	108aa	57	87aa	2.60E-14
PG23	cldA/clpB protease, Bacillus subtilis	D26185	810aa	855aa	45	855aa	7.10E-122
PG24	Putative hemolysin, Streptococcus mutans	AF051356	445aa	417aa	29	432aa	1.80E-29

Protein name	Homology description	Genbank accession number	Length of homolog protein	FASTA homology results		
				Identity %	Overlap	E value
PG25	Cysteine protease, <i>Porphyromonas gingivalis</i>	U54601	1723aa	293aa	42	1.10E-12
PG27	TonB-linked adhesin, <i>Porphyromonas gingivalis</i>	Y07818	1097aa	312aa	45	3.20E-41
PG28	Cysteine protease / hemagglutinin, <i>Porphyromonas gingivalis</i>	S75942	886aa	843aa	35	8.38aa
PG30	Putative NlpD lipoprotein, <i>Aquifex aeolicus</i>	AE0000754	187aa	337aa	42	1.42aa
PG31	Hemolysin (HvC), <i>Serpulina hydysenteriae</i>	X73141	141aa	151aa	31	1.23aa
PG32	Major outer membrane protein (ompF), <i>Pseudomonas aeruginosa</i>	M94078	350aa	391aa	26	3.82aa
PG33	Major outer membrane protein (ompF), <i>Pseudomonas fluorescens</i>	L21200	317aa	385aa	32	1.63aa
PG34	Putative membrane protein, <i>Rhodobacter capsulatus</i>	Q07396	193aa	190aa	48	1.90aa
PG35	Colicin 1 receptor, <i>Escherichia coli</i>	J04229	863aa	833aa	25	5.90aa
PG36	Outer membrane antigen (oma89), <i>Pasteurella multocida</i>	U60439	789aa	891aa	21	8.94aa
PG37	Cationic outer membrane protein (ompH), <i>Yersinia enterocolitica</i>	M34854	164aa	174aa, 170aa	27	1.68aa

Protein name	Homology description	Genbank accession number	Length of homolog protein	Length of <i>P. gingivalis</i> protein	FASTA homology results		
					Identity %	Overlap	E value
PG38	Cationic outer membrane protein (ompH), <i>Yersinia enterocolitica</i>	M33654	164aa	163aa	23	160aa	5.90E-05
PG39	Outer membrane protein (susC) <i>Bacteroides thetaiotaomicron</i>	L49338	1038aa	827aa	24	347aa	1.50E-06
PG40	Heme receptor (Hut A), <i>Vibrio cholera</i>	Q86444	693aa	772aa	23	722aa	4.90E-09
PG41	Outer membrane protein (tolC), <i>Escherichia coli</i>	X54049	496aa	462aa	22	436aa	4.60E-09
PG42	Neuramindase, <i>Micromonospora viridis faciens</i>	D01045	647aa	492aa	32	375aa	2.10E-22
PG43	Immunoreactive outer membrane protein (omp28), <i>Brucella melitensis</i>	U30815	250aa	245aa	24	178aa	0.0015
PG44	Macrophage infectivity potentiator, <i>Legionella israelensis</i>	U92208	242aa	276aa	35	218aa	9.10E-18
PG45	Outer membrane protein, <i>Neisseria meningitidis</i>	AFO21245	787aa	775aa	21	609aa	0.0034
PG46	Outer membrane protein 85, <i>Neisseria gonorrhoeae</i>	UB1859	792aa	774aa	31	117aa	0.00098
PG47	Outer membrane protein (susC) <i>Bacteroides thetaiotaomicron</i>	L49338	1038aa	807aa	20	902aa	1.00E-03

Protein name	Homology description	Genbank accession number	Length of homolog protein	Length of <i>P. gingivalis</i> protein	FASTA homology results		
					Identity %	Overlap	E value
PG48	Immunoglobulin binding surface protein (slr22), <i>Streptococcus pyogenes</i>	X75750	305aa	431aa	25	269aa	5.20E-05
PG49	Flmbrillin (orf2), <i>Porphyromonas gingivalis</i>	D42087	453aa	330aa	23	298aa	0.062
PG50	Outer membrane protein (susC) <i>Bacteroides thetaiotaomicron</i>	L49338	1038aa	840aa	26	570aa	1.00E-11
PG51	PGaA antigen, <i>Porphyromonas gingivalis</i>	X95938	202aa	202aa	54	126aa	1.20E-25
PG52	Alkaline protease secretion apparatus (aptF) <i>Pseudomonas aeruginosa</i>	X64558	481aa	455aa	21	427aa	3.50E-06
PG53	Protein export protein (tolC), <i>Salmonella enteritidis</i>	U25178	491aa	444aa	23	430aa	6.20E-11
PG54	Protease I, <i>Achromobacter lyticus</i>	I5128	653aa	940aa	24	695aa	1.50E-22
PG55	Flmbrillin (orf3), <i>Porphyromonas gingivalis</i>	D42087	670aa	670aa	43	686aa	4.90E-108
PG56	Cysteine protease <i>Porphyromonas gingivalis</i>	U68468	364aa	1202aa	1274aa	212aa	0.00012
PG57	Cysteine protease, <i>Porphyromonas gingivalis</i>	U68468	1358aa	924aa	922aa, 921aa	742aa	1.40E-23
PG60	Outer membrane protein 11, <i>Helicobacter pylori</i>	AE000562	168aa	547aa	25	183aa	2.20E+00
PG61	Ferric pseudobactin M114 receptor protein (pbuA), <i>Pseudomonas sp.</i>	X73412	826aa	749aa	22	585aa	1.00E-05

29

Protein name	Homology description	Genbank accession number	Length of homolog protein	Length of <i>P. gingivalis</i> protein	FASTA homology results		
					Identity %	Overlap	E value
PG66	Attachment and invasion protein (ail), <i>Salmonella typhimurium</i>	AF007380	105aa	206aa	21	140aa	1.90E+00
PG68	Serum opacity factor, <i>Streptococcus pyogenes</i>	U02280	1015aa	1224aa	24	170aa	2.10E-01
PG69	Vacuolating cytotoxin (vacA), <i>Helicobacter pylori</i>	UB3261	180aa	425aa	32	111aa	1.20E+00
PG70	Outer membrane protein, <i>Neisseria gonorrhoea</i>	U52069	117aa	266aa	22	153aa	6.90E+00
PG71	Gliding motility protein (glmA), <i>Flavobacterium johnsoniae</i>	AF007381	578aa	834aa	23	572aa	3.90E-25
PG75	Class 3 outer membrane porin (porB), <i>Neisseria meningitidis</i>	U07191	332aa	391aa	23	239aa	4.80E-01
PG81	Outer membrane protein (ompA), <i>Shigella dysenteriae</i>	V01344	351aa	>235aa	26	186aa	3.10E-01
PG82	Outer membrane protein (alkD), <i>Pseudomonas oleovorans</i>	X05936	230aa	434aa	26	136aa	2.80E+00
PG83	Gliding motility protein (glmA), <i>Flavobacterium johnsoniae</i>	AF007381	578aa	926aa	21	639aa	8.50E-09
PG87	Hypothetical protein, <i>Mycobacterium tuberculosis</i>	AI021942	877aa	781aa	29	784aa	2.20E-34

Protein name	Homology description	Genbank accession number	Length of homolog protein	Length of <i>P. gingivalis</i> protein			FASTA homology results		
				Identity %	Overlap	E value	Identity %	Overlap	E value
PG89	NADH-ubiquinone oxidoreductase, Helicobacter pylori	AE000631	512aa	259aa	24	168aa	3.90E-01		
PG91	Neuraminidase (nanH), <i>Bacteroides fragilis</i>	D20493	544aa	540aa	24	251aa	1.60E+00		
PG92	Hypothetical protein, <i>Mycobacterium tuberculosis</i>	AL021942	877aa	771aa	29	770aa	8.00E-30		
PG93	Cytodherence protein P1, <i>Mycoplasma pneumoniae</i>	X02191	219aa	270aa	41	63aa	6.90E-01		
PG94	Arginyl endopeptidase, <i>Porphyrromonas gingivalis</i>	D20470	991aa	1157aa	24	320aa	7.60E-08		
PG95	Sensor protein (EVGS), <i>Escherichia coli</i>	D14008	1197aa	901aa	28	511aa	2.60E-17		
PG105	Plasma cell membrane glycoprotein, Human	P22413	873aa	449aa	34	404aa	5.60E-33		
PG106	Hypothetical secreted protein, <i>Helicobacter pylori</i>	O24951	242aa	248aa	30	252aa	7.80E-22		
PG107	Cell division ATP binding protein, <i>Mycobacterium leprae</i>	O32883	229aa	248aa, 241aa, 232aa	46	193aa	1.20E-26		
PG108	ABC transporter, <i>Archaeoglobus fulgidis</i>	O29244	228aa	219aa	51	219aa	3.80E-41		
PG109	Proteinase IV, <i>Escherichia coli</i>	F84936	618aa	505aa, 589aa	38	597aa	1.10E-57		
PG110	Protein translocase, <i>Staphylococcus aureus</i>	O08446	843aa	513aa	43	521aa	6.00E-71		
PG111	ABC transporter, <i>Synechocystis</i> sp.	P73758	574aa	>720aa	40	570aa	1.70E-73		
PG112	Glycosyl transferase, <i>Erwinia amylovora</i>	Q48034	351aa	375aa, 362aa	31	363aa	1.60E-32		

31

Protein name	Homology description	Genbank accession number	Length of homolog protein	Length of <i>P. gingivalis</i> protein	FASTA homology results
					Identity % Overlap E value
PG113	Heat shock protein (dnak), <i>Treponema pallidum</i>	AB001203	635aa	640aa	62 64aa 9.10E-138
PG114	Dihydrolipamide dehydrogenase, <i>Clostridium magnum</i>	Q59299	578aa	449aa	37 45aa 3.80E-84
PG115	Zinc protease, <i>Escherichia coli</i>	P31928	931aa	941aa	27 890aa 6.60E-57
PG116	Heat shock protein (HtpG), <i>Escherichia coli</i>	P10213	624aa	681aa	32 627aa 4.60E-48
PG117	Transcriptional regulator, <i>Aquifex aeolicus</i>	CB86591	506aa	481aa	39 389aa 2.40E-49
PG118	ABC transporter, <i>Bacillus subtilis</i>	H70019	261aa	250aa	59 251aa 1.50E-60
PG119	ATP-dependent protease, <i>Aquifex aeolicus</i>	CB86827	444aa	461aa	46 458aa 1.80E-77
PG120	Nitrogen assimilation regulatory protein, <i>Bradyrhizobium sp.</i>	P10578	480aa	457aa	49 242aa 3.80E-45
PG121	Cobalamin synthesis protein, <i>Bacillus megaterium</i>	E1331323	387aa	602aa	36 224aa 9.20E-37
PG122	Outer membrane integrity (lola), <i>Haemophilus influenzae</i>	P71397	819aa	443aa	37 441aa 1.90E-54
PG123	Fimbillin, <i>Porphyromonas gingivalis</i>	D1034032	490aa	479aa	32 180aa 7.30E-48
PG124	Heat shock protein (dns), <i>Leptospira interrogans</i>	AF007813	360aa	383aa	46 356aa 2.30E-57

WO 99/29870

PCT/AU98/01023

32

Protein name	Homology description	Genbank accession number	Length of homolog protein	FASTA homology results		
				Identity %	Overlap	E value
PG125	Cobalamin biosynthesis protein(CBIK), <i>Salmonella typhimurium</i>	Q05592	264aa	283aa	37	259aa 3.70E-26
PG126	ABC-type permease, <i>Pseudomonas aeruginosa</i>	O88878	326aa	356aa	33	333aa 1.30E-30
PG127	Endonuclease excision repair protein (uvrB), <i>Pseudomonas aeruginosa</i>	X83486	670aa	678aa	58	675aa 1.10E-134

Table 3: Results of PSORT, SignalP and TopPred analysis of the proteins. The signal present column indicates the presence of a signal sequence detected with either PSORT or SignalP. The terms in parentheses indicates the type of signal sequence as determined by PSORT. The cell location & probability values are generated by PSORT and represent the probability of the protein being in the cell compartments outer membrane (OM), inner membrane (IM), periplasmic space (PC) or cytoplasm (C). The number of transmembrane domains (TMDs) was determined by TopPred and does not include uncleavable signal sequences.

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMDs
							OM	IM	PS		
PC1_386	451aa	Y		1	24	34	0	0	0.22	N	0
PC2_424	1012aa	Y		1	20	20	0.94	0	0.33	R	3
PC2_425	1014aa	Y		2	17	17	0.94	0	0.29	F	3
PC3_434	223aa	Y (lipoprotein)		1	-	18	0.79	0.78	0	K	3
PC4_447	672aa	Y (lipoprotein)		1	22	22	0.79	0.7	0	R	0
PC5_458	215aa	Y		1	40	35	0	0.25	0	R	0
PC6_475	324aa	N		1	-	-	0	0	0.2	S	1
PC7_487	104aa	N		1	7	0	0	0.12	0	E	3
PC8_498	598aa	N		1	-	-	0	0	0.22	N	0
PC9_499	550aa	N		2	-	-	0	0	0.35	N	0

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid of TMD's
							OM	IM	PS	
PG8	500	458aa	N	3	-	-	0	0	0.34	N
PG8	501	326aa	N	4	-	-	0	0	0.24	N
PG9	515	1268aa	N	1	7	-	0	0	0.22	E
PG9	516	1232aa	N	2	-	-	0	0	0.39	E
PG9	517	1174aa	N	3	-	-	0	0	0.47	E
PG10	387	185aa	N	1	-	-	0	0	0.11	K
PG11	400	313aa	Y	1	22	26	0.24	0	0.33	R
PG12	411	271aa	Y (lipoprotein)	3	27	29	0.79	0.7	0	R
PG13	419	157aa	Y	1	23	25	0.94	0	0.29	0
PG14	420	331aa	Y (uncharable)	1	35	28	0	0.58	0	K
PG15	421	287aa	Y	2	24	16	0	0.11	0	K
PG16	422	569aa	Y (lipoprotein)	1	24	16	0.79	0.7	0	G
PG18	423	681aa	Y	1	30	0	0.58	0	0	K
PG21	426	321aa	Y	2	24	27	0.34	0	0.37	G
PG22	427	108aa	Y (uncharable)	1	41	11	0	0.26	0	P
PG23	428	459aa	N	1	-	-	0	0.12	0	A

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMDs
							OM	IM	PS		
PG24	429	417aa	Y	1	19	0	0.44	0	0	N	3
PG25	430	293aa	Y	1	27	28	0.2	0	0.82	R	0
PG27	431	312aa	N	1	-	0	0	0	0.28	Q	1
PG28	432	843aa	Y	1	21	21	0.93	0	0.24	O	1
PG29	433	280aa	Y	1	18	16	0.28	0	0.94	H	1
PG30	435	137aa	Y	1	21	21	0.24	0	0.4	K	1
PG31	436	151aa	N	1	-	0	0	0	0.3	T	0
PG32	437	391aa	Y	1	20	20	0.62	0	0.13	K	0
PG33	438	386aa	Y	1	26	26	0.81	0	0.31	E	1
PG34	439	190aa	Y	1	-	13	0	0.5	0	A	5
PG34	440	186aa	Y (uncleavable)	2	-	47	0	0.5	0	A	4
PG35	441	833aa	Y	1	22	22	0.94	0	0.37	F	1
PG36	442	891aa	Y (uncleavable)	1	-	40	0	0.31	0	F	2
PG37	443	174aa	Y (uncleavable)	1	28	24	0	0.35	0	K	0

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMDs	
							OM	IM	PS			
PG37	444	170aa	Y (uncleavable)	2	24	20	0	0.35	0	K	0	
PG38	445	163aa	Y	1	18	0.21	0	0.93	0	K	1	
PG39	446	327aa	Y	1	26	0.03	0	0.25	0	R	3	
PG40	448	272aa	Y	2	19	0.04	0	0.32	0	F	4	
PG41	449	362aa	Y	2	27	0.25	0	0.54	0	Q	2	
PG42	450	392aa	Y	5	30	0	0	0.00	0.13	Q	2	
PG43	451	245aa	Y (uncleavable)	2	28	22	0	0.38	0	K	1	
PG44	452	276aa	Y	1	19	24	0.15	0	0.88	0	K	0
PG45	453	275aa	Y (lipoprotein)	1	19	23	0.79	0.7	0	P	4	
PG46	454	274aa	Y	1	27	27	0.73	0	0.22	0	F	2
PG47	455	367aa	Y	1	24	0.94	0	0.38	0	P	2	
PG48	456	431aa	Y	1	24	0	0.1	0	0	R	1	
PG49	457	333aa	Y (uncleavable)	1	24	18	0	0.12	0	I	0	
PG50	458	348aa	Y	1	21	0.94	0	0.34	0	F	3	
PG51	460	202aa	Y	1	26	25	0.2	0	0.61	S	0	
PG52	461	455aa	Y (uncleavable)	1	23	21	0	0.18	0	R	1	

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMD's
							OM	IM	PS	C	
PG53	462	444aa	Y	1	14	17	0.36	0	0.22	D	2
PG54	463	940aa	Y	1	27	20	0.86	0	0.25	Q	5
PG55	464	870aa	Y (lipoprotein)	1	23	23	0.79	0.7	0	K	2
PG56	465	1282aa	Y (uncharavable)	1	-	21	0	0.04	0	K	4
PG58	466	1274aa	N	2	-	-	0	0	0	C	
PG57	467	925aa	Y	1	28	24	0.53	0	0.2	C, D, K	5
PG57	468	922aa	Y	2	25	21	0.53	0	0.2	P	3
PG57	469	921aa	Y	3	24	20	0.53	0	0.2	P	3
PG58	470	583aa	Y	1	24	24	0.82	0	0.19	0	F
PG58	471	588aa	Y	2	20	20	0.82	0	0.19	0	P
PG59	472	349aa	Y	1	37	-	0	0.18	0	F	1
PG59	473	345aa	Y	2	36	56	0.92	0	0.15	0	F
PG59	474	330aa	Y	3	21	41	0.93	0	0.25	0	P
PG60	476	547aa	Y	1	28	28	0.93	0	0.25	0	F
PG61	477	749aa	Y	2	21	21	0.94	0	0.29	0	F
PG62	478	494aa	Y	1	21	21	0.93	0	0.24	0	P

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid of TMD's
							OM	IM	PS	
PC33	479	294aa	Y	1	20	20	0.93	0	0.24	P
PC34	480	204aa	Y	1	20	20	0.93	0	0.19	P
PC35	481	243aa	Y	1	18	18	0.93	0	0.25	P
PC36	482	208aa	Y	1	21	21	0.94	0	0.3	F
PC37	483	650aa	Y	1	28	38	0.93	0	0.27	Y
PC38	484	1226aa	Y	1	25	25	0.91	0	0.31	Y
PC38	485	1225aa	Y	2	24	24	0.91	0	0.31	Y
PC39	486	125aa	Y	1	29	29	0.93	0	0.21	P
PC70	488	260aa	Y	1	18	24	0.99	0	0.24	P
PC71	489	934aa	Y	2	20	30	0.94	0	0.21	N
PC72	490	399aa	Y	1	27	27	0.94	0	0.32	H
PC73	491	382aa	Y	2	20	20	0.94	0	0.3	L
PC74	492	222aa	Y	1	24	24	0.94	0	0.32	L
PC75	493	391aa	Y	1	26	28	0.94	0	0.3	H
PC76	494	446aa	Y	1	21	22	0.94	0	0.32	V
PC77	495	308aa	Y	2	28	28	0.94	0	0.38	K

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMDs
							OM	IM	PS		
PG78	496	314aa	Y	1	23	23	0.94	0	0.29	D	0
PG79	497	285aa	Y	1	-	32	0.93	0	0.26	C	2
PG80	502	240aa	Y	1	19	19	0.93	0	0.22	0	2
PG81	366	>235aa	Y	1	26	20	0.93	0	0.21	0	1
PG82	503	434aa	Y	1	30	24	0.93	0	0.2	Q	1
PG83	504	926aa	Y	1	23	57	0.93	0	0.21	S	1
PG84	505	400aa	Y	1	25	25	0.93	0	0.25	N	1
PG84	506	398aa	Y	2	23	23	0.93	0	0.25	0	1
PG85	507	581aa	Y	1	20	20	0.93	0	0.46	L	2
PG86	508	239aa	Y	1	44	-	0	0	0.12	H	0
PG86	509	211aa	Y	2	16	46	0.91	0	0.03	H	0
PG87	510	781aa	Y	1	26	47	0.89	0	0.21	O	2
PG88	511	271aa	Y	2	28	19	0.89	0	0.25	P	0
PG88	512	270aa	Y	3	27	18	0.89	0	0.25	P	0
PG88	513	287aa	Y	4	24	15	0.89	0	0.23	P	0
PG89	514	259aa	Y	2	23	25	0.88	0	0.35	N	1

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMD's
							OM	IM	PS		
PG90	518	229aa	Y	1	22	21	0.85	0	0.44	K	0
PG90	519	228aa	Y	2	21	20	0.85	0	0.44	K	0
PG91	520	540aa	Y	1	25	25	0.85	0	0.30	E	0
PG92	521	771aa	Y	2	19	19	0.85	0	0.3	R	3
PG93	522	776aa	Y	1	25	25	0.85	0	0.37	R	4
PG94	523	1157aa	Y	1	23	28	0.8	0	0.25	Q	5
PG95	524	881aa	Y (ipoprotein)	1	-	19	0.79	0.87	0	V	1
PG96	525	563aa	Y	1	23	23	0.40	0	0.33	K	0
PG97	526	437aa	Y	1	23	23	0.32	0	0.85	O	0
PG98	527	318aa	Y (ipoprotein)	1	19	19	0.79	0.7	0	L	1
PG99	528	461aa	Y (uncleavable)	1	22	20	0	0	0.3	R	0
PG100	388	279aa	Y	1	20	18	0.26	0	0.54	I	0
PG101	268	>157aa	N (ORF incomplete)	-	-	-	-	-	-	R	1
PG102	389	562aa	Y	1	29	29	0.19	0	0.4	S	3
PG102	390	558aa	Y	2	25	25	0.26	0	0.46	S	3
PG104	391	381aa	Y	1	17	17	0.62	0	0.22	R	0

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMD's
							OM	IM	PS		
PG105	392	449aa	Y	1	22	19	0.31	0	0.91	0	P
PG106	393	246aa	Y	2	41	49	0	0	0	0.02	L
PG107	394	246aa	N	1	-	-	0	0	0	0.32	D
PG107	395	241aa	N	2	-	-	0	0	0	0.3	D
PG107	396	232aa	N	3	-	-	0	0	0	0.21	D
PG108	397	219aa	N	1	-	-	0	0	0	0.19	R
PG109	398	595aa	Y	1	35	37	0.26	0	0.93	0	Y
PG109	399	589aa	Y	2	29	31	0.27	0	0.93	0	Y
PG110	401	>523aa	N	1	-	-	0	0	0	0.38	Incomplete
PG111	278	>720aa	N (ORF incomplete)	-	-	-	-	-	-	-	0
PG112	402	375aa	Y	1	-	43	0	0.12	0	0	C
PG112	403	362aa	Y	2	-	30	0	0	0.12	0	N
PG113	404	640aa	N	1	-	-	0	0	0	0.25	K
PG114	405	449aa	N	1	-	-	0	0.12	0	0	G
PG115	406	941aa	Y	1	23	22	0.13	0	0.92	0	Q
PG116	407	884aa	N	1	-	-	0	0.12	0	0	L

Protein name	Protein seqID number	Protein Length	Signal Present	Methionine in ORF	SignalP cleavage site	PSORT cleavage site	Cell Location & probability			C-terminal Amino Acid	Number of TMD's
							OM	IM	PS		
PG117	408	464aa	N	1	-	-	0	0.19	0	L	1
PG118	409	250aa	N	1	-	-	0	0	0.27	E	1
PG119	410	461aa	N	1	-	-	0	0.28	0	E	2
PG120	412	457aa	N	1	-	-	0	0	0.21	E	0
PG121	413	602aa	N	1	-	-	0	0	0.31	E	3
PG122	414	443aa	N	1	-	-	0	0	0.14	Q	4
PG123	415	479aa	Y	2	22	22	0.26	0	0.04	O	0
PG124	416	383aa	N	1	-	-	0	0	0.23	D	2
PG125	417	283aa	Y	1	23	15	0.16	0	0.93	R	1
PG126	418	356aa	N	1	-	-	0	0.52	0	D	9
PG127	532	578aa	N	1	-	-	0	0	0.28	A	2

**Table 4:** Percentage identity and percentage similarity of various proteins with the 70 amino acids from the C-terminal of the *P. gingivalis* arginine protease 1 (RGP1), arginine protease 2 (RGP2), and the cysteine protease/hemagglutinin (prtT).

5

Protein name	Percent identity			Percent similarity		
	RGP1	RGP2	prtT	RGP1	RGP2	prtT
PG21	17	29	21	40	57	49
PG25	43	41	9	64	73	14
PG27	41	33	7	73	74	11
PG28	21	26	34	49	57	74
PG54	19	13	16	40	43	33
PG57	11	14	19	20	24	34
PG91	31	21	39	57	53	74
PG96	0	13	20	0	24	43
PG97	10	26	33	14	47	61
PG98	16	20	0	47	54	0
PG99	19	0	26	41	0	54
PG100	20	21	24	39	57	41
PG101	11	16	27	17	39	60
PG102	27	20	31	50	61	61
PG104	16	23	26	46	44	49

**Table 5:** Percentage identity and percentage similarity of various proteins with the TonBIII box of *P. gingivalis*.

Protein name	Percent identity	Percent similarity
PG2	46	71
PG13	57	93
PG35	50	96
PG47	39	71
PG50	54	93

**Cloning, expression and purification of recombinant *P. gingivalis* genes.**

**PG1**

- 5        Oligonucleotides to the 5' and 3' regions of the deduced protein were used to amplify the gene of interest from a preparation of *P. gingivalis* W50 genomic DNA using the TaqPlus Precision PCR System (Stratagene) and a PTC-100 (MJ Research) thermal cycler or similar device. The 5' oligonucleotide primer sequence was GCGCCATATGCTGCCGAACCGGCC,
- 10      the 3' oligonucleotide primer sequence was GCGCCTCGAGTCAATTCAATTCTTATAGAG. The PCR fragment was purified, digested with Nde I, Xho I restriction enzymes (Promega) and ligated into the corresponding sites of the plasmid pProEx-1 (Gibco-BRL) and transformed into *E. coli* ER1793 cells (a gift from Elizabeth Raleigh, New
- 15      England Biolabs). A resulting clone expressing the correct insert was selected and induced with or without 0.1mM IPTG (Promega) for expression of the recombinant protein. Expression of the recombinant protein was determined by SDS-PAGE analysis and Western Blot using the one of the rabbit antisera described above or an anti-hexahistidine antibody (Clontech)
- 20      that detects the hexahistidine tag that was fused to the *P. gingivalis* recombinant protein. PG1 was purified by disruption of the *E. coli* cells by sonication in binding buffer (Novagen) and solubilisation by the addition of sarkosyl (N-Lauroyl sarcosine) to a 1% final concentration. There after the preparation was diluted to 0.1% sarkosyl in binding buffer, bound to a
- 25      Nickel-nitrilotriacetic acid column (Ni-NTA; Qiagen), after washing bound proteins were eluted with 1M imidazole in elution buffer (Novagen) according to the Qiagen recommendations with 0.1% sarkosyl added to all buffers. Following purification samples were dialysed against 500mM NaCl, 20mM Tris, 0.1% sarkosyl at pH7.4 to remove the imidazole, concentrated as
- 30      required and stored at 4°C until used. Purity and antigenicity were assessed by SDS-PAGE and Western blot using selected antisera (from those described above) and the protein concentration was determined by the BCA assay (Pierce).

**PG2**

The methods used for PG2 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACATGAAAAGAACGACGC, the 3' oligonucleotide primer sequence was CGCGAGATCTGAAAGACAACGTGAATACC and the PCR product was cloned into pGex-stop RBS(IV) (Patent application WO9619496, JC Cox, SE Edwards, I Frazer and EA Webb. Variants of human papilloma virus antigens) using the BstZ 171 and Bgl II restriction sites. 2% sarkosyl was used to solubilise PG2 and 8M urea was added to the solubilisation buffer and to all other buffers. Urea was removed from the purified protein by sequential dialysis (4M then 2M then 1M then 0.5M then 0M urea all in 50mM Tris, 500mM NaCl, 0.1% sarkosyl, pH7.4). Purified protein was stored at 4°C until required.

**15 PG3**

The methods used for PG3 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGTATACATGAAGAAATCAAGTGTAG, the 3' oligonucleotide primer sequence was GCGCAGATCTCTCAGCGTACCTTGCTGTG and DNA was amplified with Pfu DNA polymerase (Stratagene). The PCR product was cloned directly into pCR-Blunt and transformed into *E. coli* Top10F'(InVitrogen) before subcloning into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* BL21DE3 (Pharmacia Biotech). The following modifications were made to the purification of PG3 from the PG1 method. Cells expressing the recombinant protein were disrupted by sonication in binding buffer and the insoluble inclusion bodies concentrated by centrifugation. Inclusion bodies were then solubilised in 6M urea (Sigma) in binding buffer and eluted with 6M urea added to the elution buffer. In some instances 6M guanidine hydrochloride (Sigma) was used instead of urea for these steps. Urea (or guanidine hydrochloride when it was substituted) was removed from the purified protein by sequential dialysis against reducing levels of urea (3M then 1.5M then 0.5M then 0M urea all in 50mM Tris, 500mM NaCl, 8% glycerol, pH7.4). Purified protein was stored frozen at -80°C until required. Protein concentration was determined by the Coomassie Plus protein assay (Pierce).

**PG4**

The methods used for PG4 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CTTCTGTATACTTACAGCGGACATCATAAAATC, the 3' oligonucleotide 5 primer sequence was TTCCAGGAGGGTACCACGCAACTCTTCTTCGAT and DNA was amplified with the Tth XL PCR kit (Perkin Elmer). The PCR product was cloned into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Kpn I restriction sites and transformed into *E. coli* ER1793.

**10 PG5**

The methods used for PG5 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was TTGCAACATATGATCAGAACGATACTTTCA, the 3' oligonucleotide primer sequence was AGCAATCTCGAGCGGTTCATGAGGCCAAAGC and DNA was 15 amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24 (Novagen) using the Nde I and Xho I restriction sites and transformed into *E. coli* BL21 (Pharmacia Biotech). Removal of urea was not proceeded past 1M urea as the protein was insoluble at lower concentrations of urea. Purified protein was stored at 4°C until required.

20

**PG6**

The methods used for PG6 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was TAAACATATGTGCCTCGAACCCATAATTGCTCCG, the 3' oligonucleotide 25 primer sequence was CGTCCGCGGAAGCTTGATCGGCCATTGCTACT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Nde I and Hind III restriction sites and transformed into *E. coli* BL21.

**30 PG8**

The methods used for PG8 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACTGGAGTTCAAGATTGTG, the 3' oligonucleotide primer sequence was CGCGAGATCTGTTTCTGAAAGCTTTC and DNA was 35 amplified with the TaqPlus Precision PCR System. The PCR product was

cloned into the expression plasmid pProEx-1 using the Nde I and Xho I restriction sites and transformed into *E. coli* ER1793.

#### PG8A

5 PG8A is a shortened version of PG8 and has the first 173 amino acids removed. The methods used for PG8A were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACATGGAAAACCTTAAAGAAC, the 3' oligonucleotide primer sequence was CGCGAGATCTGTTTCTGAAAGCTTTTC and DNA was  
10 amplified with the TaqPlus Precision PCR System. The PCR product was cloned into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* ER1793. Prior to dialysis of the purified protein EDTA (Sigma) was added to a final concentration of 10mM.

15

#### PG10

The methods used for PG10 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGATATCATGGATAAAAGTGAGCTATGC, the 3' oligonucleotide primer sequence was CGCGAGATCTTTGTTGATACTCAATAATTG and DNA was  
20 amplified with the TaqPlus Precision PCR System. The PCR product was digested with Eco RV and Bgl II and ligated into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* ER1793.

25

#### PG11

The methods used for PG11 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGTATACATGAGAGCAAACATTGGCAGATACTTCCG, the 3'  
30 oligonucleotide primer sequence was GCGCAGATCTGCGCAAGCGCAGTATATGCC and DNA was amplified with Tli DNA polymerase (Promega). The PCR product was cloned into pCR-Blunt and transformed into *E. coli* Top10F' before subcloning into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and  
35 transformed into *E. coli* ER1793. PG11 was purified by solubilisation of *E. coli* cells with 2% sarkosyl in binding buffer (Qiagen) which was diluted to

0.1% sarkosyl in binding buffer, bound to a Nickel-nitrilotriacetic acid column (Ni-NTA; Qiagen), after washing bound proteins were eluted with 1M imidazole (0.7% CHAPS (Sigma) in elution buffer; Qiagen) according to the Qiagen recommendations. Following purification samples were dialysed  
5 against 500mM NaCl, 20mM Tris, 0.7% CHAPS, 20% glycerol (Sigma) at pH7.4 to remove the imidazole, concentrated as required and stored at 4°C until used.

**PG12**

10 The methods used for PG12 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGTATACTGAATAGCAGACATCTGACAATCACAATCATTGCCGG, the 3' oligonucleotide primer sequence was GCGCAGATCTGCTGTTGTGAGTGCAGTTGTTAAGTG and DNA was  
15 amplified with Tli DNA polymerase. The PCR product was cloned into pCR-Blunt and transformed into *E. coli* Top10F'cells before subcloning into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* BL21. Purification of the recombinant protein was essentially the same as PG11 except 0.5% DHPC  
20 (1,2-Diheptanoyl-sn-glycero-3-phosphocholine; Avanti) in 50mM Tris, 50mM NaCl, pH8.0 was used to solubilise the inclusion bodies instead of sarkosyl and the DHPC was diluted to 0.1% before addition to the Ni-NTA and 0.1% DHPC was added to all buffers.

**PG13**

The methods used for PG13 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCCATATGCGGACAAAAACTATCTTTTTGCG, the 3' oligonucleotide primer sequence was  
30 GCGCTCGAGGTTGTAATCGAATCGCTATTCGAGC and DNA was amplified with Tli DNA polymerase. The PCR product was cloned the expression plasmid pET24b using the Nde I and Xho I restriction sites and transformed into *E. coli* BL21. Purification of the recombinant protein was essentially the same as PG3 using 6M urea and 1% NOG (n-octyl glucoside; Sigma) was added to the dialysis buffer. Removal of urea was not proceeded  
35

past 2M urea as the protein was insoluble at lower concentrations of urea. Purified protein was stored at 4°C until required.

#### PG14

- 5       The methods used for PG12 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGGCCATGACGGACAACAAACAGTAATATCG, the 3' oligonucleotide primer sequence was GCGCCTCGAGTTACTTGCATGATCACGGACATAACC and DNA was  
10      amplified with Tli DNA polymerase. The PCR product was cloned the expression plasmid pProEx-1 using the Ehe I and Xho I restriction sites and transformed into *E. coli* BL21. Purification of the recombinant protein was essentially the same as PG12.

#### 15      PG15

- The methods used for PG15 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CAAAAGTATACTAATAATATCATTCTCAA, the 3' oligonucleotide primer sequence was GCTTATGGTACCTTGGTCTATCTATTAT and DNA was  
20      amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Kpn I restriction sites and transformed into *E. coli* ER1793.

#### PG22

- 25      The methods used for PG22 were essentially the same as for PG1 with the following exceptions. The 5' oligonucleotide primer sequence was CCCCGGATCCGATGCGACTGATCAAGGC, the 3' oligonucleotide primer sequence was CCCCCTCGAGCGAACGGGGTCATAGCC and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was  
30      cloned into the expression plasmid pET24b using the Bam HI and Xho I restriction sites and transformed into *E. coli* BL21DE3. Once PG22 was purified dialysis was performed in the same manner as for PG1 but in the presence of 1M imidazole.

**PG24**

- The methods used for PG24 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGGTATACATGAATTACCTGTACATAC, the 3' oligonucleotide primer sequence was CGCGGGATCCGTTGATTGGTCGTCGATGG and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was digested with Bst Z171 and Bam HI and ligated into the expression plasmid pGex-stop RBS(IV) using the Bst Z171 and Bgl II restriction sites and transformed into *E. coli* ER1793. Due to the low level of expression of PG24 purification was not proceeded with except on small scale.

**PG24A**

- A modified version of PG24 was also cloned and expressed. PG24A is the same as PG24 with the predicted N-terminal sequence removed. The methods used for PG24A were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was CGCGCATATGGAGATTGCTTCTTCTTCG, the 3' oligonucleotide primer sequence was CGCGCTCGAGTTAGTTGATTGGTCGTCG and DNA was amplified with the TaqPlus Precision PCR System. The PCR product was cloned into the expression plasmid pProEx-1 using the Nde I and Xho I restriction sites and transformed into *E. coli* ER1793. Purification of the recombinant protein was essentially the same as PG3 except 8M urea was used to solubilise the inclusion bodies and in the buffers used for the Ni-NTA column purification. Urea was removed by sequential dialysis (4M then 2M, then 1M then 0.5M then 0M urea all in 50mM Tris, 500mM NaCl, 8% glycerol, pH7.4). Purified protein was stored frozen at -80°C until required.

**PG29**

- The methods used for PG29 were essentially the same as for PG3 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGATATCGCTAGCATGAAAAAGCTATTCTC, the 3' oligonucleotide primer sequence was GCGCAGATCTCTCGAGTTGCCATCGGATTGCGGATTG and DNA was amplified with Pfu DNA polymerase being used. The PCR product was cloned into pCR-Blunt (InVitrogen) and transformed into *E. coli* Top10F before subcloning into the expression plasmid pGex-stop RBS(IV)

using the EcoR V and Bgl II restriction sites and transformed into *E. coli* BL21. 6M urea was used throughout the purification process.

#### PG30

- 5       The methods used for PG30 were essentially the same as for PG3 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TACGGAATTCTGTGACCCCCGTCAAGAAATGTGCGC, the 3' oligonucleotide primer sequence was
- 10      CTATCGGGCCGCTTGATCCTCAAGGCTTGCCCCGG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates of PG30. 10ml cultures of
- 15      recombinant *E. coli* were grown to an OD of 2.0 ( $A_{600\text{nm}}$ ) in terrific broth and the cells were induced with 0.5mM IPTG and samples taken for analysis at 4 hours post induction. Purification was not done for these studies.

#### PG31

- 20      The methods used for PG31 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was CGGGGAATTCGCAAAATCAATTCTATGCTGAA, the 3' oligonucleotide primer sequence was CTATCGGGCCGCTGTATGCAATAGGGAAAGCTCCGA and DNA was amplified with the Tth XL PCR kit. The PCR product was
- 25      cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG32

- 30      The methods used for PG32 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGCAGAATTCCAGGAGAATACTGTACCGGCAACG, the 3' oligonucleotide primer sequence was
- 35      CTATCGGGCCGCTTGGAGCGAACGATTACAACAC and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
5 for these studies.

#### PG33

The methods used for PG33 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was  
10 removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCAGAATTCCAAGAACGACTACACAGAACAAA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTTCCGCTGCAGTCATTACTACAA and DNA was amplified with the Tth XL PCR kit. The PCR product was  
15 cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG35

The methods used for PG35 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was GCGCGAATTGATGAAACAACTAAACATTATCAGC, the 3' oligonucleotide primer sequence was GCGTGCGGCCGCGAAATTGATCTTGTACCGACGA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.  
30

#### PG36

The methods used for PG36 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was AAAGGAATTCTACAAAAAGATTATTGCCGTAGCA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCGAACTCCTGTCCGAGCACAAAGT and DNA was amplified with the Tth XL PCR kit. The PCR product was  
35

cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

5

#### PG37

The methods used for PG37 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was TGGCGAATTCAAACGGTTTGATTTGATCGGC, the 3' oligonucleotide primer sequence was CTATCGGCCGCCTGCTAAAGCCCATCTTGCTCAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10

15

#### PG38

The methods used for PG38 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CCTCGAATTCCAAAAGGTGGCAGTGGTAAACACT, the 3' oligonucleotide primer sequence was CTATCGGCCGCCTGATTCCGAGTTTCGCTTTAC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

20

25

#### PG39

The methods used for PG39 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCTGGATCCCAAGGGTCAAGGTATCGGGCTAT, the 3' oligonucleotide primer sequence was CTATCGGCCGCAGATTGACGAGGAGACGCAGGT and DNA was

30

35

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
5 for these studies.

#### PG40

The methods used for PG40 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was  
10 removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCTGAATTCAAGACGGACAACGTCCGACAGAT, the 3' oligonucleotide primer sequence was  
CTATGCCGCCCGCGAAGTTGACCATAACCTTACCCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
15 for these studies.

#### PG41

The methods used for PG41 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GACTGAATTCCAAAACGCCTCCGAAACGACGGTA, the 3'  
25 oligonucleotide primer sequence was  
CTATGCCGCCCGCTTGTGGAAATCCCCATGCCGTT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
30 studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG42

The methods used for PG42 were essentially the same as for PG30  
35 with the following exceptions. The 5' oligonucleotide primer sequence was GTTGAAATTGCAAATAACTCTTTGGCGAAG, the 3' oligonucleotide

primer sequence was  
GAGTGC GGCCGCTTGGACATCGAAGAGATCGTC and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
expression plasmid pET24a using the Eco RI and Not I restriction sites and  
5 transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
studies were carried out on whole *E. coli* lysates. Purification was not done  
for these studies.

**PG43**

10 The methods used for PG43 were essentially the same as for PG30  
with the following exceptions. The 5' oligonucleotide primer sequence was  
GCGCGAATTCAAAAAAGAAAAACTTGCGATTGCG, the 3' oligonucleotide  
primer sequence was CTATGC GGCGCCTCAAAGCGAAAGAACGCTTAAC  
and DNA was amplified with the Tth XL PCR kit. The PCR product was  
15 cloned into the expression plasmid pET24a using the Eco RI and Not I  
restriction sites and transformed into *E. coli* BL21DE3. Expression studies  
and immunoreactivity studies were carried out on whole *E. coli* lysates.  
Purification was not done for these studies.

**PG44**

20 The methods used for PG44 were essentially the same as for PG30  
with the following exceptions. The predicted N-terminal signal sequence was  
removed from the recombinant protein. The 5' oligonucleotide primer  
sequence was AGCCGAATTCTGTAAAGAAAAATGCTGACACTACC, the 3'  
25 oligonucleotide primer sequence was  
CTATGC GGCGCCTTTCCGGGCTTGATCCCCAT and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
expression plasmid pET24a using the Eco RI and Not I restriction sites and  
transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
30 studies were carried out on whole *E. coli* lysates. Purification was not done  
for these studies.

**PG45**

35 The methods used for PG45 were essentially the same as for PG30  
with the following exceptions. The predicted N-terminal signal sequence was  
removed from the recombinant protein. The 5' oligonucleotide primer

sequence was GACAGGATCCTGCTCCACCACAAAGAACATCTGCCG, the 3' oligonucleotide primer sequence was

CTATGCGGCCGCGAAGGGATAGCCGACAGCCAAAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the

- 5 expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10 PG46

The methods used for PG46 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CTCGGAATTCCGTTATGTGCCGGACGGTAGCAGA, the 3'

- 15 oligonucleotide primer sequence was CTATGCGGCCGCGAACGGATAGCCTACTGCAATGT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity

- 20 studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG47

- 25 The methods used for PG47 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was

- CGCCGAATTCCAAACAGTGGTGACCGGTAAAGGTGATCGATTAGAA, the 3' oligonucleotide primer sequence was

- 30 CTATGCGGCCGCGAACGTITACACGAATAACGGTAGACCAAGTGCAGCC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates.

- 35 Purification was not done for these studies.

**PG48**

The methods used for PG48 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAAAAATCCAAGCAGGTACAGCGA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTCGTAACCATAGTCTGGGTTTG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG49**

15 The methods used for PG49 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GAACGGATCCAACGAGCCGGTGGAAAGACAGATCC, the 3' oligonucleotide primer sequence was  
20 GAGTGCAGCCGCTAACATCTCGACTTCATACTTGTACCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
25 for these studies.

**PG50**

30 The methods used for PG50 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GCTGGGATCCGCGACAGACACTGAGTTCAAGTAC, the 3' oligonucleotide primer sequence was CTATGCGGCCGCGAACCTCACTACCAAGCCCATGT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG51

- 5       The methods used for PG51 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TCTTGAATT CGCGCAAAGTCTTT CAGCACCGAA, the 3' oligonucleotide primer sequence was
- 10      CTATGCGGCCGC ACTTT CGTGGGATCACTCTCTT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done
- 15      for these studies.

#### PG52

- 20      The methods used for PG52 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was AGAAGAATTCAAACGGACAATCCTCCTGACGGCA, the 3' oligonucleotide primer sequence was CTATGCGGCCGCGAAGTCTTGCCCTGATAGAAATC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG53

- 30      The methods used for PG53 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCTGAATT CGCGAATCCCCTACGGGCCAATCG, the 3' oligonucleotide primer sequence was
- 35      CTATGCGGCCGCGTCCGAAAGGCAGCCGTAA TAGG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and

transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

5 PG54

The methods used for PG54 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGCTGAATTCCAGATTCGTTGGAGGGAAACCC, the 3' 10 oligonucleotide primer sequence was CTATGCGGCCGCTGCTTACGATCTTGCTCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity 15 studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG55

The methods used for PG55 were essentially the same as for PG30 20 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGAGGGATCCGAGCTCTATTGCGATGGCGAG, the 3' oligonucleotide primer sequence was GAGTGCCTGGCTTACCTGACTTGTACGAAT and DNA was 25 amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

30

PG56

The methods used for PG56 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was 35 AAATGGATCCCCAAAAATTGAGCTTTGATG, the 3' oligonucleotide primer sequence was CTATGCGGCCGCTTGATTGTAATTTCGGTATC and DNA was amplified with the Tth XL PCR kit. The PCR product was

cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

5

**PG57**

The methods used for PG57 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGGATCCCAAGAGATCTCAGGCATGAATGCA, the 3' oligonucleotide primer sequence was GAGTGGCGGCCGCTCGGCCTCTTATCTCTACCTTTTC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10

The methods used for PG58 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGGTGAATTCCAACCCCCACGAAATACAGAAACC, the 3' oligonucleotide primer sequence was

15

GAGTGGCGCCGCTGAAAGTCCAGCTAAAACCGGGCGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done

20

25

for these studies.

30

**PG58**

The methods used for PG59 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAACAAGAGAAGCAGGTGTTCAT, the 3'

oligonucleotide primer sequence was  
GAGTGC GGCCGCTGAAGATGCTCTTATCGTCAAACG and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
expression plasmid pET24a using the Eco RI and Not I restriction sites and  
5 transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
studies were carried out on whole *E. coli* lysates. Purification was not done  
for these studies.

**PG60**

10 The methods used for PG60 were essentially the same as for PG30  
with the following exceptions. The predicted N-terminal signal sequence was  
removed from the recombinant protein. The 5' oligonucleotide primer  
sequence was GGCGGAATTCCAGATGCTCAATACTCCTTCGAG, the 3'  
oligonucleotide primer sequence was  
15 GAGTGC GGCCGCTGAAGAGGTAGGAGATATTGCAGAT and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
expression plasmid pET24a using the Eco RI and Not I restriction sites and  
transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
studies were carried out on whole *E. coli* lysates. Purification was not done  
20 for these studies.

**PG61**

The methods used for PG61 were essentially the same as for PG30  
with the following exceptions. The predicted N-terminal signal sequence was  
25 removed from the recombinant protein. The 5' oligonucleotide primer  
sequence was AGCAGAATTCCCCGTCTCAAACAGCGAGATAGAT, the 3'  
oligonucleotide primer sequence was  
GAGTGC GGCCGCTGAAATCGATTGTCAGACTACCCAG and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
30 expression plasmid pET24a using the Eco RI and Not I restriction sites and  
transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
studies were carried out on whole *E. coli* lysates. Purification was not done  
for these studies.

**PG62**

The methods used for PG62 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer 5 sequence was TGCTGAATTCCAGCGGTTCCGATGGTGCAGGGA, the 3' oligonucleotide primer sequence was GAGTGCAGGCCGCTGAAGTGAAATCCGACACGCAGCTG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and 10 transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG63**

15 The methods used for PG63 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAAGAACGAAACTGCATCTGAC, the 3' oligonucleotide primer sequence was 20 GAGTGCAGGCCGCTGAAGTGACGAAACACCCACGCC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done 25 for these studies.

**PG64**

The methods used for PG64 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was 30 removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGAGTCGTCCTGCTCTTAGACTG, the 3' oligonucleotide primer sequence was GAGTGCAGGCCGCTGAAGCGAACACCGAGACCCACAAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the 35 expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG65

- 5       The methods used for PG65 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGGATCCATCGGACAAAGCCGCCGGCACTT, the 3' oligonucleotide primer sequence was
- 10      GAGTGCGGCCGCTAAAGCGGTAAACCTATGCCACGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done
- 15      for these studies.

#### PG66

- The methods used for PG66 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GTTTGAATTCCAAGACGTTATCAGACCATGGTCA, the 3' oligonucleotide primer sequence was
- 20      GAGTGCAGCCGCTAAATGAGTGGAGAGCGTGGCCAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done
- 25      for these studies.

#### PG67

- The methods used for PG67 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GAACGAGCTCGCGAACGTCCTATGGCCGGAGCA, the 3' oligonucleotide primer sequence was
- 30      GAGTGCAGCCGCTATAACCAAGTATTCTGATGGGACG and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Sac I and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
5 for these studies.

#### PG68

The methods used for PG68 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was  
10 removed from the recombinant protein. The 5' oligonucleotide primer sequence was GCTTGGCGCCGCCCTATGAAAGATTGCAGAT, the 3' oligonucleotide primer sequence was GGTGCTCGAGTATACTCAACAAGCACCTTATGCAC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Not I and Xho I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
15 for these studies.

#### PG69

The methods used for PG69 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGGAAGGGGAGGGGAGTGCCGA, the 3'  
25 oligonucleotide primer sequence was GAGTGCAGCCGCTGAAGCTGTAGCGGGCTTGAACCA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
30 for these studies.

#### PG70

The methods used for PG70 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer  
35

sequence was CGGTGGATCCTCGCAAATGCTCTTCAGAGAAT, the 3' oligonucleotide primer sequence was  
5 GAGTGC GGCCGCTAAACGAAATATCGATAACCAACATC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10 PG71

The methods used for PG71 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGAACAAATACCCTCGATGTACAC, the 3'  
15 oligonucleotide primer sequence was GAGTGC GGCCGCTATTGCCGGTAGGATTCCCTTGTC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
20 studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

PG72

The methods used for PG72 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTGGAGAGCGACTGGAGACGGACAGC, the 3'  
25 oligonucleotide primer sequence was GAGTGC GGCCGCTATGATTGCCTTCAGAAAAGCTAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed TGCTGAATTGGAGAGCGACTGGAGACGGACAGC into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out  
30 on whole *E. coli* lysates. Purification was not done for these studies.

**PG73**

The methods used for PG73 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGGTGAATTCCAACAGACAGGACCGGCCAACGC, the 3' oligonucleotide primer sequence was GAGTGCAGCCGCTTAAGAAAGGTATCTGATAGATCAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG74**

15 The methods used for PG74 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAAGAAAATAATACAGAAAAGTCA, the 3' oligonucleotide primer sequence was  
20 GAGTGCAGCCGCTGAGGTATAATCCTATGCCAATACT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
25 for these studies.

**PG75**

30 The methods used for PG75 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGGGATCCGCTCAGGAGCAACTGAATGTGGTA, the 3' oligonucleotide primer sequence was  
35 GAGTGCAGCCGCTGTGGAACAAATTGCGCAATCCATC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG76

- 5       The methods used for PG76 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCAGAATTGGAAACGCCACAGAGCTTTGGGAA, the 3' oligonucleotide primer sequence was
- 10      GAGTGCAGCCCGCTTACCTGCACCTTATGACTGAATAc and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done
- 15      for these studies.

#### PG77

- 20      The methods used for PG77 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAAGAGAAAAAGGATAGTCTCTCT, the 3' oligonucleotide primer sequence was
- 25      GAGTGCAGCCCGCTTCTTATGCCATAGAACAGG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done
- for these studies.

#### PG78

- 30      The methods used for PG78 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAGGATTCTCCCACGGTAGCAAT, the 3' oligonucleotide primer sequence was
- 35      GAGTGCAGCCCGCTATGATAGTAAAGACTGGTTCT and DNA was

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
5 for these studies.

#### PG79

The methods used for PG79 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was  
10 TGCTGAATTCTCGTAGTGACGCTGCTCGTAATTGTC, the 3' oligonucleotide primer sequence was  
GAGTGCAGGCCGCTGCCGTCCCTGCCTTCTGCCTGACG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
15 studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG80

20 The methods used for PG80 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAATTCCAAAACGTGCAGTTGCAC TACGAT, the 3' oligonucleotide primer sequence was  
25 GAGTGCAGGCCGCTGTTGAAAGTCCATTGACCGCAAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
30 for these studies.

#### PG81

The methods used for PG81 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was  
35 removed from the recombinant protein. The 5' oligonucleotide primer sequence was GTTTGAATTCCAGGATTTCTATGAAATAGGA, the 3'

oligonucleotide primer sequence was  
GAGTGCGGCCGCTTGTATTACAAAAAGTCTTACG and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
expression plasmid pET24a using the Eco RI and Not I restriction sites and  
5 transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
studies were carried out on whole *E. coli* lysates. Purification was not done  
for these studies.

**PG82**

10 The methods used for PG82 were essentially the same as for PG30  
with the following exceptions. The predicted N-terminal signal sequence was  
removed from the recombinant protein. The 5' oligonucleotide primer  
sequence was GAACGAATTCCAGAACAAACTTACCGAGTCG, the 3'  
oligonucleotide primer sequence was  
15 GAGTGCGGCCGCTGTTCAAGCTTTAACCA and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
expression plasmid pET24a using the Eco RI and Not I restriction sites and  
transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
studies were carried out on whole *E. coli* lysates. Purification was not done  
20 for these studies.

**PG84**

25 The methods used for PG84 were essentially the same as for PG30  
with the following exceptions. The predicted N-terminal signal sequence was  
removed from the recombinant protein. The 5' oligonucleotide primer  
sequence was TGCTGGATCCCAGAATGATGACATCTCGAAGAT, the 3'  
oligonucleotide primer sequence was  
GAGTGCGGCCGCTATTGCGTCCCCGGCCACTACGTCC and DNA was  
amplified with the Tth XL PCR kit. The PCR product was cloned into the  
30 expression plasmid pET24a using the Bam HI and Not I restriction sites and  
transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
studies were carried out on whole *E. coli* lysates. Purification was not done  
for these studies.

**PG85**

The methods used for PG85 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was CGGTGAATTCGTACCAACGGACAGCACGGAATCG, the 3' oligonucleotide primer sequence was GAGTGCAGCCGCTCAGATTGGTGCTATAAGAAAGGTA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG86**

15 The methods used for PG86 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGGATCCAAACGCATGATCATCTCATCGAA, the 3' oligonucleotide primer sequence was  
20 GAGTGCAGCCGCTGTGGTTCAGGCCGTGGCAAATCT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
25 for these studies.

**PG87**

The methods used for PG87 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGGAATTCCAGAGCTATGTGGACTACGTCGAT, the 3' oligonucleotide primer sequence was  
30 GAGTGCAGCCGCTATTACTGTGATTAGCGCGACGCTG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
35

WO 99/29870

PCT/AU98/01023

studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG88**

- 5 The methods used for PG88 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCAGAATTGCCGAATCGAAGTCTGTCTTTTC, the 3' oligonucleotide primer sequence was
- 10 GAGTGCAGCCGCTCGCAAGTAACGCTTAGTGGGA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.
- 15

**PG89**

- 20 The methods used for PG89 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAATCGAAGTTAAAGATCAAGAGC, the 3' oligonucleotide primer sequence was
- GAGTGCAGCCGCTATTAGTCCAAAGACCCACGGTAAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.
- 25

**PG90**

- 30 The methods used for PG90 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAAACAAACGACGAACAGTAGCCGG, the 3' oligonucleotide primer sequence was
- GAGTGCAGCCGCTTTTGTTGATACTGTTGGGC and DNA was
- 35

WO 99/29870

PCT/AU98/01023

72

amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
5 for these studies.

**PG91**

The methods used for PG91 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was  
10 removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCCAGACGATGGGAGGAGATGTC, the 3' oligonucleotide primer sequence was  
GAGTGCAGCGCTTCCACGATGAGCTCTACGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the  
15 expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG92**

The methods used for PG92 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGAATTCGCCGATGCACAAAGCTCTGTCT, the 3'  
25 oligonucleotide primer sequence was  
GAGTGCAGCGCTTCCAGAGCATTGCTTAGTCGTA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
30 studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG93**

The methods used for PG93 were essentially the same as for PG30  
35 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer

sequence was GGCCGAGCTCCAAGAGGAAGGTATTTGGAATACC, the 3' oligonucleotide primer sequence was  
5 GAGTGCAGGCCGCTGCGAATCACTGCGAAGCGAATTAG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Sac I and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10 **PG94**

The methods used for PG94 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAGCTCCAAGAGGAAGGTATTTGGAATACC, the 3'  
15 oligonucleotide primer sequence was GAGTGCAGGCCGCTTGTCTACGATCATTCTT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity  
20 studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG95**

The methods used for PG95 were essentially the same as for PG30  
25 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAGCTCTGTGGAAAAAAAGAAAAACACTCT, the 3' oligonucleotide primer sequence was  
30 GAGTGCAGGCCGCTAACGTCTCCTGTGCGCTCCCCGG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Sac I and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG96**

The methods used for PG96 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAGCTCCAAACGCAAATGCAAGCAGACCGA, the 3' oligonucleotide primer sequence was GAGTGCAGGCCGCTTTGAGAATTTCATTGTCTCACG and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Sac I and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

**PG97**

15 The methods used for PG97 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCGGGATCCCAGTTGTTCCGGCTCCCACCACA, the 3' oligonucleotide primer sequence was  
20 GAGTGCAGGCCGCTCTGTTGATGAGCTTAGTGGTATA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done  
25 for these studies.

**PG98**

The methods used for PG98 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was AGCAGAATTCCAAGAAAGAGTCGATGAAAAAGTA, the 3' oligonucleotide primer sequence was  
30 GAGTGCAGGCCGCTAGCTGTAAACATTAAGTTTTATTGAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and

immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

#### PG99

- 5       The methods used for PG99 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was TGCTGAATTCAAGGACAATTCTTCTTACAAACCT, the 3' oligonucleotide primer sequence was
- 10      GAGTGCAGGCCGCTTCGAATCACGACTTTCTCACAAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done
- 15      for these studies.

#### PG100

- 20      The methods used for PG100 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCAGAATTCCAGTCTTGAGCACAATCAAAGTA, the 3' oligonucleotide primer sequence was
- 25      GAGTGCAGGCCGCTGATGCCAGCTTGATGCTCTTAGC and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done
- 30      for these studies.

#### PG101

- 35      The methods used for PG101 were essentially the same as for PG30 with the following exceptions. The 5' oligonucleotide primer sequence was TGCTGAATTCAAAGGCAAGGGCGATCTGGTCGGG, the 3' oligonucleotide primer sequence was
- 40      GAGTGCAGGCCGCTCTCTCGAACCTGGCCGAGTA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the

expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

5

**PG102**

The methods used for PG102 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GGCCGAATTCCAGATGGATATTGGTGGAGACGAT, the 3' oligonucleotide primer sequence was GAGTGCAGCGCTCTACAATGATTTTCACGAA and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Eco RI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

10

**PG104**

The methods used for PG104 were essentially the same as for PG30 with the following exceptions. The predicted N-terminal signal sequence was removed from the recombinant protein. The 5' oligonucleotide primer sequence was GAACGGATCCAACGTGTCTGCTCAGTCACCCGA, the 3' oligonucleotide primer sequence was GAGTGCAGCGCTCTGAGCGATACTTTGCACGTAT and DNA was amplified with the Tth XL PCR kit. The PCR product was cloned into the expression plasmid pET24a using the Bam HI and Not I restriction sites and transformed into *E. coli* BL21DE3. Expression studies and immunoreactivity studies were carried out on whole *E. coli* lysates. Purification was not done for these studies.

15

20

25

30

**Animal antisera and human patient sera.**

Various antisera were raised for detecting the expression and refolding of the recombinant *P. gingivalis* proteins. A whole cell antisera was raised by injecting New Zealand White rabbits with 3 doses of sonicated

*P. gingivalis* (strain W50) containing approximately 2mg of protein. The first dose was given in Freunds complete adjuvant (FCA) and the second and third doses were given in Freunds incomplete adjuvant (IFA) at 3 week intervals.

- Doses (1ml) were given intramuscularly into the hind legs and rabbits bled 7 days after the last dose, the blood clotted and serum removed and stored at -20°C until required. A second rabbit antisera was produced in a similar manner but using a sarkosyl insoluble fraction (each dose was 0.69mg of protein) derived from *P. gingivalis* W50 according to the method of Doidg and Trust T. et al 1994 as the immunogen. A third rabbit antisera was produced 10 in a similar manner to the first only the sarkosyl soluble fraction (1mg of protein per dose) derived from *P. gingivalis* W50 cells according to the method of Doidg P. and Trust TJ. (1994 Infect Immun 62:4526-33) was used as the immunogen.

- A "protected rat serum" pool was also used in these studies and was obtained from rats immunised with formalin killed whole *P. gingivalis* cells in FIA (strain ATCC 33277; 2 doses of  $2 \times 10^9$  cells, 3 weeks apart). Rats were then challenged 2 weeks after their last dose with live *P. gingivalis* cells (strain 33277) given orally as previously described (Klaussen B. et al. 1991, Oral Microbiol Immunol 6:193-201) and the serum obtained from these rats 6 weeks after the final challenge inoculation at the time of sacrifice.

- Human sera were obtained from adult patients undergoing treatment or assessment for periodontitis at an outpatient clinic. These patients had at least 6 teeth with 6mm attachment loss and had *P. gingivalis* present in their sub-gingival plaque as detected using a *P. gingivalis* specific DNA probe. 25 Sera was pooled from these patients and compared to a pool of sera from periodontally healthy patients.

#### Immunization and Murine Lesion Model Protocols

- 30 The mouse abscess model was used to assess the efficacy of immunising mice with recombinant *P. gingivalis* proteins in protecting mice from formation of a subcutaneous abscess. This model has been used by others as a predictor of potential vaccines against periodontal disease (Bird PS, et al. 1995 J. Periodontol. 66:351-362. BALB/c mice 6-8 weeks old were 35 immunised by subcutaneously injecting them with 0.1 ml containing either 10 or 20 $\mu$ g of recombinant *P. gingivalis* protein, 20 $\mu$ g of *E. coli* lysate protein,

2 x 10<sup>9</sup> formalin killed cells of *P. gingivalis* strain 33277 emulsified in incomplete Freund's adjuvant (IFA; Sigma) on day 0. At day 21 mice were re-injected with the same dose and then bled 1 week later and evaluated for antibody levels. At day 35 mice all mice were challenged with 5 approximately 2 x 10<sup>9</sup> cells of live *P. gingivalis* (ATCC 33277) by subcutaneous injection in the abdomen. Following challenge mice were monitored daily for weight loss and the size of the lesion measured for the next 10 days. Lesion sizes were measured by length and width and expressed as mm<sup>2</sup>. Groups were statistically analysed using a Kruskal-Wallis one-way 10 ANOVA and were also individually examined using the unpaired t test or Mann-Whitney rank sum test using the Instat statistical package.

Figure 1 shows the results of one experiment at day 4 after challenge (lesions were at maximum size at this time point). Control mice immunised with *E. coli* lysate showed large lesions while mice immunised with killed 15 cells of *P. gingivalis* strain 33277 were fully protected. This indicates that whole cells provide protection against *P. gingivalis* while *E. coli* protein immunised mice were not protected. Mice given the various PG recombinant proteins showed significant levels of protection for PG2, PG22, PG24 and PG29 ( $p < 0.05$  unpaired t test) while PG8A was not quite significantly 20 different ( $p = 0.07$ ) compared to the *E. coli* control group.

Figure 2 shows the results of a separate experiment using combinations of recombinant proteins. Mice given PG1 + PG2 showed a significant level of protection compared to control mice give *E. coli* lysate ( $p < 0.026$  unpaired t test).

25

### Immunoscreening

Cloned candidates were cultured in 15ml of Terrific broth, induced with IPTG and sampled at 4h post-induction. One ml of culture was 30 removed, pelleted and the cells resuspended in a volume of PBS determined by dividing the OD A<sub>600nm</sub> of the culture by 8. An aliquot of lysate (100μl) was added to 100μl of 2x sample reducing buffer (125mM Tris pH 6.8, 20% glycerol, 4% SDS, 80mM DTT, 0.03% bromophenol blue) and boiled for 10min. SDS-PAGE was performed according to the method of Laemmli UK. 35 1970 (Nature 227:680-685) using 4-20% 1.0mm Tris-Glycine gels (Novex) according to the manufacturers recommendations. Proteins were transferred

onto Hybond-C Extra nitrocellulose membranes (Amersham) by transblotting and the membranes were then blocked for 2h at room temperature (RT) in 5% skim milk in 20mM Tris, 0.5M NaCl, 0.05% Tween-20, pH 7.5 (TTBS).

Immunoscreening was performed separately with the rabbit

- 5 anti-*P. gingivalis* whole cell serum, the rat protective serum, a pool of human periodontal patients serum, and in many cases an anti-T7-Tag antibody HRP conjugate (Novagen). Prior to use, the rabbit, rat and human sera were diluted 1/5000, 1/1000 and 1/500 respectively in 5% skim milk in TTBS and absorbed with 100µl (for the rabbit serum) or 250µl (for the rat and human sera) *E. coli* extract (20mg/ml; Promega) for 6h at RT.

Membranes were incubated overnight at RT with the absorbed antisera, or for 1 hr at RT with 1/5000 diluted anti-T7-Tag conjugate.

- Following 3x10min washes with TTBS, HRP-conjugated anti-rabbit (Silenus), anti-mouse (Silenus) or anti-human (KPL) antibody, diluted 1/5000 in 5% 15 skim milk in TTBS, was added for 1h at RT. Membranes were washed as before, prior to addition of TMB membrane peroxidase substrate (KPL) for detection of immunoreactive proteins. Results of reactivity for the recombinant *P. gingivalis* proteins is shown in Table 7.

- In addition some of the sera (pooled sera diluted 1/1000) from the 20 mice immunised with *P. gingivalis* recombinant proteins (prior to challenge) were analysed for their reactivity against Western blots of whole native W50 *P. gingivalis* proteins using similar techniques as those outlined above. PG2, PG8A, PG29 and PG3 all showed bands at a similar molecular weight to that of the recombinant PG protein in the native W50 blot. This indicates that PG 25 proteins are expressed in the W50 strain and that the recombinant proteins have at least some identical immunogenicity to the native proteins.

#### m-RNA analysis

##### 30 Hot Phenol RNA Extraction

- P. gingivalis* W50 cells (150ml culture) were grown anaerobically to mid log phase ( $OD A_{600}=0.18$ ) mixed with 50% glycerol and stored at -70°C until RNA extraction. Cells were pelleted by centrifugation at 6000g, and 35 resuspended in 8ml ASE (20mM NaOAc, 0.5% SDS, 1mM EDTA). An equal volume of 20mM NaOAc(pH 4.5)-saturated phenol was added and mixed by

shaking for 30 seconds, incubated at 65°C for 5 minutes, followed by a further 5 second shaking and repeated incubation. After cooling, 2ml chloroform was added and mixed by shaking for 5 seconds, and the mixture spun at 10000g for 10 minutes at 4°C. The top aqueous phase was transferred and re-extracted by repeating the phenol and chloroform steps. The aqueous phase was transferred again and 100U RNase inhibitor (RNAsin; Promega) were added. RNA was precipitated with 3 volumes 100% ethanol at -20°C overnight. The RNA precipitate was recovered by centrifugation at 10000g at 4°C for 15 minutes, then washed with 100% ethanol, dried and resuspended in 600µl sterile, deionised, dH<sub>2</sub>O with 1µl of fresh RNase inhibitor. RNA was aliquoted and stored at -70°C. The RNA concentration was determined spectrophotometrically. A formaldehyde RNA gel confirmed RNA integrity (Sambrook J. et al. 1989, Molecular Cloning. A laboratory manual. Cold Spring Laboratory Press, New York. 2nd Edition).

15

#### RT-PCR

The isolated RNA was used as a template for Reverse Transcription (RT) to produce cDNA. Varying RNA concentrations were used for the RT as each RNA transcript was potentially present at different levels. Subsequent amplification of the cDNA was performed using Polymerase Chain Reaction (PCR). RT-PCR was performed using GeneAmp® RNA PCR Kit (Perkin Elmer) according to the manufacturer's protocol with the following exception to the PCR; 35 cycles were performed as follows: Melt phase 95°C for 30 seconds, Anneal phase varied between 50-60°C for 30 seconds, Extension phase 72°C for 1 minute. Amplification was performed in a PTC-100 Programmable Thermal Controller (MJ Research Inc.). As a control to demonstrate that the amplified product did not arise from contaminating DNA, Reverse Transcriptase (RTase) was omitted from a parallel tube. The PCR products were examined against DNA markers (GIBCO 1kB ladder) on a 1% agarose gel stained with ethidium bromide.

RT-PCR results are shown in Table 6 using the oligonucleotide primers as used in "Cloning, expression and purification of recombinant *P. gingivalis* genes" section described above, except for the following changes. For PG1 the 3' reverse primer used was

- GCGCCTCGAGATTCA<sup>1</sup>TTCCTTATAGAG, for PG4 the 5' forward primer was CTCTTGTCGACTACAGCGGACATCATAAAATC and the 3' reverse primer was TTCCACCTCGAGTTAACGCAACTCTTCTTCGAT, for PG6 the 5' forward primer was TAAAGAATTCTGCCTCGAACCCATAATTGCTCCG, for PG10 the 5' forward primer was CGCGATATGGATAAAAGTGAGCTATGC and the 3' reverse primer was CGCGCTCGAGTTGTGATACTCAATAATTC, for PG13 the 5' forward primer was GCCCGGCCATGCGAACAAAATCTTTTGCG and the 3' reverse primer was
- PG10      GCCCGGCCCTAGTTGAATCGAATCGCTATTGAGC. Amplification of *P. gingivalis* transcripts is a likely indication that RNA for a specific candidate is present and that the protein is produced. However, where there is no amplification achieved this does not indicate that this gene is never transcribed and may be the result of the culture conditions or the state of the cells when harvested.

Table 6. Expression of PG m-RNA with *in vitro* grown *P. gingivalis* W50. The symbols are + band visible on agarose gel, - no band present on agarose gel, ND not detected.

20

PG #	RNA μg	Annealing temp. °C	RT-PCR	PCR (-RT)	Approx. fragment size bp	Expected fragment size bp
1	0.15	55	+	-	1300	1362
2	1.0	50	+	-	3200	3051
3	0.15	60	+	-	720	690
4	2.9	55	-	-	N.D.	2000
5	0.02	50	+	-	1000	947
6	1.0	55	+	-	1000	972
8A	0.15	50	+	-	1200	1278
10	0.15	55	+	-	590	585
11	0.10	60	+	-	960	942
12	0.02	60	+	-	880	831
13	1.0	50	+	-	2150	2274
14	0.15	60	+	-	1050	996

PG #	RNA μg	Annealing temp. °C	RT-PCR	PCR (-RT)	Approx. fragment size bp	Expected fragment size bp
22	1.0	60	-	-	N.D.	228
24	1.0	55	+	+	1150	1194
29	0.15	60	+	-	880	885

5 Table 7: Immunoblot results of proteins expressed in *E.coli* against rabbit, rat  
10 and human antisera. Deduced MW was calculated from amino acid sequence  
of the *P. gingivalis* proteins, some of which had their N-terminal signal  
sequences removed. Apparent MW was determined from SDS-PAGE gels.  
The N- and C-terminal tags add approximately 2.5 KDa to the deduced MW of  
the recombinant proteins. The symbols are + positive, - negative, +/- weak  
positive, ND not done.

Protein number	Deduced MW (KDa)	Apparent MW (KDa)	Antisera reactivity			
			T7	Rabbit	Rat	Human
PG1	47.5	63	ND	-	-	-
PG2	112.4	125.7	ND	+	-	-
PG3	22.6	18.3	ND	- <sup>a</sup>	-	-
PG4	75	90.6	ND	-	-	-
PG5	34.9	43.8	ND	-	-	-
PG6	36.7	47.1	ND	-	-	-
PG8	67.5	63.1	ND	- <sup>b</sup>	-	-
PG8A	47.7	90.6	ND	-	-	-
PG10	21.3	25.5	ND	+	-	+
PG11	36.2	42.4	ND	-	-	-
PG12	30.7	30.6	ND	-	-	-
PG13	84.5	101	ND	-	-	-
PG14	36	42.4	ND	-	+	+
PG22	8.6	11.1	ND	-	-	-
PG24A	47	63.1	ND	-	-	-
PG29	31.1	40.9	ND	+	+	+

Protein number	Deduced MW (kDa)	Apparent MW (kDa)	Antisera reactivity			
			T7	Rabbit	Rat	Human
PG30	35.1	46.9	+	-	-	-
PG31	16.7	-	-	-	-	-
PG32	41.2	59.5	+	+	+	-
PG33	39.9	52.7	+	+	+	-
PG35	92.6	116.6	+	-	-	-
PG36	98.9	120.2	-	-	-	-
PG37	18.8	23.1	+	+	-	-
PG38	16.1	22.0	+	-	-	-
PG39	87.9	116.6	+	-	-	-
PG40	76.6	103.1	+	-	-	-
PG41	48.3	81.1	+	-	+	+
PG42	59.3	73.9	+	-	-	-
PG43	27.1	50.3	+	-	-	-
PG44	28.6	32.3	+	-	+	-
PG45	84	100.6	+	-	-	-
PG46	83	97.7	+	-	-	-
PG47	93.7	42.5	+	+	-	+
PG48	45.2	37.9	+	-	-	-
PG49	33.3	64.1	+	-	+	-
PG50	91.9	113.2	+	+	-	-
PG51	19.6	27.2	+	-	-	-
PG52	50.4	64.4	+	+	-	+
PG53	47.4	45.4	+	-	-	+
PG54	101.4	46.7	+	+	-	-
PG55	70.4	68.4	+	-	-	-
PG56	142.3	-	-	-	-	-
PG57	100	134.5	+	+	+	+
PG58	63	82.9	+	-	-	-
PG59	33.3	43.6	+	-	-	-
PG60	55.6	77.8	+	-	-	-
PG61	81.5	107.3	+	-	-	-

Protein number	Deduced MW (kDa)	Apparent MW (kDa)	Antisera reactivity			
			T7	Rabbit	Rat	Human
PG62	51.9	58.4	+	-	-	-
PG63	29.6	43.6	+	-	-	-
PG64	18.5	26.9	+	-	-	-
PG65	25.9	28.8	+	-	-	-
PG66	22.2	25.1	+	+	-	-
PG67	103.7	105	+	-	-	-
PG68	133.3	30.7	+	-	+	+
PG69	44.4	50.8	+	-	-	-
PG70	25.9	30.8	+	-	-	-
PG71	88.9	105.5	+	-	-	-
PG72	40.7	49.8	+	-	-	-
PG73	40.7	29	+/-	-	-	-
PG74	22.2	32.5	+	-	-	-
PG75	40.7	46.7	+	-	-	-
PG76	48.1	55.6	+	-	-	+
PG77	29.6	36.9	+	-	-	-
PG78	33.3	35.4	+	-	-	-
PG79	33.3	-	-	-	-	-
PG80	25.9	20.5	+	-	-	-
PG81	23	25.8	+	-	-	-
PG82	44.8	48.5	+	-	-	-
PG84	41.7	52.4	+	-	-	+/-
PG85	62.7	72.4	+	-	-	-
PG86	21.7	27.4	+	-	-	+/-
PG87	83	91.3	+	-	-	+
PG88	27	40.1	+	-	-	-
PG89	26.2	29.4	+	-	-	-
PG90	23	28.4	+/-	-	-	-
PG91	57.2	85.7	+	+	+	+
PG92	83.6	110.4	+	-	-	+
PG93	83.4	110.4	+	-	-	+

Protein number	Deduced MW (kDa)	Apparent MW (kDa)	Antisera reactivity			
			T7	Rabbit	Rat	Human
PG96	59.3	70.3	+	+	+	+
PG97	44.4	57.5	+	-	+	+
PG98	33.3	36	+	-	-	-
PG99	40.7	55.6	+	-	+	+
PG100	29.6	10.8	+	-	-	-
PG101	14.8	19.7, 14.1	+	-	-	-
PG102	59.3	70.3	+	-	-	+
PG104	40.7	57.5	+	-	-	+

- a. Positive reaction detected with the rabbit antiserum to sarkosyl insoluble *P. gingivalis* antigen.
- 5 b. Purified protein demonstrated weak positive reaction with the rabbit antiserum to whole *P. gingivalis*.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

**References.**

1. Lipman DJ, Pearson WR. 1985. Rapid and sensitive protein similarity searches. *Science* 277:1435-1441.
- 5 2. Horton, P. and Nakai, K. (1996). A probabilistic classification system for predicting the cellular localization sites of proteins. *Intellig. Syst. Mol. Biol.* 4: 109-115.
- 10 3. Nakai K, Kanehisa M. 1991. Expert systems for predicting protein localization sites in Gram-negative bacteria. *Proteins: Structure, Function, and Genetics* 11:95-110.
- 15 4. Nielsen H, Engelbrecht J, Brunak S and von Heijne G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* 10, 1-6.
5. Claros MG and G von Heijne. (1994). TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* 10: 685-686.
- 15 6. Borodovsky M, Rudd KE, and EV Koonin. (1994). Intrinsic and extrinsic approaches for detecting genes in a bacterial genome. *Nucleic Acids Res.* 22:4756-4767.
- 20 7. Struyve M, Moons M, Tommassen J. 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J. Mol. Biol.* 218:141-148.
8. Aduse-Opoku J, Slaney JM, Rangarajan M, Muir J, Young KA, Curtis MA. 1997. The Tla receptor protein of *Porphyromonas gingivalis* W50: a homolog of the RI precursor (PrpRI) is an outer membrane receptor required for growth on low levels of hemin. *J. Bacteriol.* 179:4778-4788.
- 25 9. Needleman SB, Munsch CD. 1970. A general method applicable to the search of similarity in the amino acid sequence of two proteins. *J. Molec. Biol.* 48: 443-453.
- 30